

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 3, 2002, 10:35:17 ; Search time 2060 Seconds

(without alignments)  
11250.360 Million cell updates/sec

Title: US-09-908-988B-1

Sequence: 1 aaagagtgtagacagagtggtt.....ataaagactcaagtgccc 1431

Scoring table: OLIGO-MNC

Gapop 60.0 , Gapext 60.0

Searched: 16154066 seqs, 8097743376 residues

Word size: 15

Total number of hits satisfying chosen parameters: 34152

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database: EST

1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estnu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_hic:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_hic:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: gb\_gss:\*  
18: em\_gss\_hum:\*  
19: em\_gss\_inv:\*  
20: em\_gss\_pln:\*  
21: em\_gss\_vit:\*  
22: em\_gss\_fun:\*  
23: em\_gss\_man:\*  
24: em\_gss\_mus:\*  
25: em\_gss\_other:\*  
26: em\_gss\_pro:\*  
27: em\_gss\_rtd:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	482	33.7	485	AK019655	AK019655 Mus muscu
2	267	18.7	461	AV006036	AV006036 AV006036
3	258	18.0	432	BB732046	BB732046 BB732046
4	168	11.7	957	AK016235	AK016235 Mus muscu
5	147	10.3	418	A2507637	A2507637 IM0349D21
6	108	7.5	316	BB140247	BB140247 BB140247

7	79	5.5	195	10	EB135648	EB135648
8	75	5.2	392	9	AV006120	AV006120 AV006120
9	68	4.8	378	10	AA918753	AA918753 EST350057
10	65	4.5	456	10	BE099923	BE099923 UI-R-BJ1-
11	65	4.5	504	9	AT172700	AT172700 UI-R-AF1-
12	65	4.5	608	12	BF284900	BF284900 EST449491
13	65	4.5	638	9	AA800245	AA800245 EST189742
14	53	3.7	449	17	A2731346	A2731346 RPI-24-1
15	52	3.6	184	10	AV361069	AV361069 AV361069
16	48	3.4	440	12	BG376557	BG376557 UI-R-CSO-
17	47	3.3	441	12	BG379943	BG379943 UI-R-CSO-
18	46	3.2	250	9	AV268013	AV268013 AV268013
19	46	3.2	386	10	BE095943	BE095943 UI-R-BU0-
20	46	3.2	467	13	BM087315	BM087315 220024 MA
21	46	3.2	483	12	BF073506	BF073506 220224 MA
22	44	3.1	512	12	BF855618	BF855618 RCO-FN020
23	44	3.1	555	12	BF855618	BF855618 RCO-FN020
24	39	2.7	339	9	AA998888	AA998888 UI-R-CO-h
25	38	2.7	319	12	BG381122	BG381122 UI-R-CSO-
26	38	2.7	329	9	AT1763576	AT1763576 UI-R-YO-a
27	36	2.5	270	9	AJ403135	AJ403135 AJ403135
28	36	2.5	354	9	AA182748	AA182748 2P40F10.S
29	36	2.5	420	9	AA447961	AA447961 zV83908.S
30	36	2.5	565	12	BF828781	BF828781 MR2-HN003
31	36	2.5	572	12	BF828778	BF828778 MR2-HN003
32	36	2.5	586	12	BF826755	BF826755 MR2-HN003
33	36	2.5	598	12	BF825744	BF825744 MR2-HN003
34	35	2.4	181	10	BB014893	BB014893 BB014893
35	35	2.4	373	10	AA483766	AA483766 56372 MAR
36	35	2.4	516	13	BM030114	BM030114 488634 MA
37	32	2.2	291	12	BF554230	BF554230 UI-R-CO-h
38	32	2.2	482	12	BF825323	BF825323 MR2-HN003
39	31	2.2	475	31	AA428229	AA428229 zV83906.r
40	31	2.2	478	14	BQ359416	BQ359416 MR2-HN003
41	31	2.2	580	12	BF828699	BF828699 MR2-HN003
42	30	2.1	415	12	BF828153	BF828153 MR2-HN003
43	30	2.1	637	14	BM890066	BM890066 fY64604.Y
44	29	2.0	553	9	AA808889	AA808889 nW16A05.S
45	29	2.0	595	13	BJ071661	BJ071661 BJ071661

## ALIGNMENTS

RESULT 1	AK019655	485 bp	musculus adult male testis cDNA, RIKEN full-length enriched library, clone:4930486E09, ring finger protein 30, full insert
LOCUS	AK019655		
DEFINITION	AK019655		
ACCESSION	AK019655		
VERSION	AK019655.1	GI:12859964	
KEYWORDS	HTC: CAP trapper.		
SOURCE	Mus musculus (strain: C57BL/6J) adult male testis cDNA to mRNA, clone:lib:RIKEN full-length enriched mouse cDNA library		
ORGANISM	Mus musculus		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
AUTHORS	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Kono, H., Okazaki, T., Muramatsu, M., and Hayashizaki, Y.		
TITLE	High-efficiency full-length cDNA cloning		
JOURNAL	Meth. Enzymol. 303, 19-44 (1999)		
MEDLINE	99279253		
PUBMED	10349636		
REFERENCE	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Kono, H., Okazaki, T., Muramatsu, M., and Hayashizaki, Y.		
AUTHORS	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes		
TITLE	Genome Res. 10 (10), 1617-1630 (2000)		
JOURNAL	20499374		
MEDLINE	11042159		
PUBMED			





ACCESSION	AV006036
VERSION	AV006036.1
KEYWORDS	GI:4782886
SOURCE	EST.
ORGANISM	house mouse. Mus musculus.
REFERENCE	Eukaryota; Metazoa; Chordata; Cranita; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 461)
AUTHORS	Cernicchi, P., Shibata, K., Ozawa, Y., Konno, H., Itoh, M., Aizawa, K., Akahira, S., Akiyama, J., Fukuda, S., Fukunishi, Y., Funayama, T., Hara, A., Hayatsu, N., Horii, F., Ishikawa, T., Itoh, M., Izawa, M., Kawai, J., Kikuchi, N., Kojima, Y., Matsuyama, T., Nitsuma, H., Oda, H., Owa, C., Sato, K., Shibata, Y., Suzuki, H., Tatenou, M., Tomatsu, Y., Tomiada, N., Watanabe, S., Yagame, M., Yamamura, T., Yokota, T., Yoshino, M., Muramatsu, M., Okazaki, Y. and Hayashizaki, Y.
TITLE	RIKEN Mouse ESTs
JOURNAL	Unpublished (1999)
COMMENT	Contact: Chie Owa Genome Science Laboratory

Email: genome-res@tc.riken.go.jp  
Thermolabile and thermostable of thermolabile enzymes by  
trehalose and its application for the synthesis of full length cDNA  
(Proc. Natl. Acad. Sci. U.S.A. 95(2):520-524 (1998))  
Transcriptional sequencing: A method for DNA sequencing using RNA  
polymerase (Proc. Natl. Acad. Sci. U.S.A. 95(7):3455-3460 (1998))  
Please visit our web site (<http://genome.ttc.riken.go.jp>) for  
further details.

FEATURES	Location/Qualifiers
source	1. .461

```

/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone_1fb="1020004119"
/clone_1fb="Mus musculus C57BL/6J heart"
/tissue_type="heart"
/note="Organ: mammary gland; Vector: p7T73D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - Oligo(dT) primer [5',
TGTACCACATCTGACATGGGAGCGGCCGACATGCTTTTTTTTTTTTTTTTTTTT
T 3']"; double-stranded cDNA was ligated to Eco RI
adaptors (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of the modified p7T73 vector.
RNA provided by Dr. Minoru Ko, Wayne State Univ. library
constructed and normalized by Bento Soares and M.Fatima
Bonaldo."

```

Query Match	18.7%;	Score 267;	DB 9;	Length 461;
Best Local Similarity	99.7%;	Pred. No. 9.9e-117;		
Matches 317;	Conservative	0;	Mismatches 1;	Indels 0;
				Gaps 0.

Accession	Sequence	Length
OY	TATGAGAGCATGAGCAATTCTCTGTGAGCGTGGAGACACTGGCCGAAATGTTGCGAAC	1173
Db	TATGAGAGCATGAGCAATTCTCTGTGAGCGTGGAGACACTGGCCGAAATGTTGCGAAC	203
OY	ATGCACTTCCAGCGCGCGCGCTGGGATGAAGAGATGACGATGCGCTTTTGGATGGG	12333
Db	ATGCACTTCCAGCGCGCGCGCTGGGATGAAGAGATGACGATGCGCTTTTGGATGGG	263

QY	1234	GAGGAGGCGCAATCCGGGCGCTGGAGAGAGACGGCTGCACCTGCCCCAAGAAGCTCAGGCGCTG	1293
Db	264	GAGGAGGCGCAATCCGGGCGCTGGAGAGAGACGGCTGCACCTGCACCAAGAAGCTCAGGCGCTG	323
QY	1294	CACATGACCCGACTCTATATCCAGAGCGCACACACCCGAAACGGGAGGCAAGGATGCTGTAAGA	135
Db	324	CACATGACCCGACTCTATATCCAGAGCGCACACCCGAAACGGGAGGCAAGGATGCTGTAAGA	383

QY	1414	AAAGACTCAAGTGTCC	1431
QY	1354	TTCTGGCAGAGACACCGGGCCACCAAGTCTGGCTTCCGGCCCGGGGAAGTTTCAT	1413
Db	384	TTCTGGCAGAGACACCGGGCCACCAAGTCTGGCTTCCGGCCCGGGGAAGTTTCAT	443
QY	1414	AAAGACTCAAGTGTCC	1431
Db	444	AAAGACTCAAGTGTCC	461

RESULT	3
B732046	
LOCUS	432 bp mRNA linear EST 12-Oct-2001
DEFINITION	B732046 RIKEN full-length enriched 12 days embryo whole body Mus musculus cDNA clone E970020A01 3', mRNA sequence.
ACCESSION	B732046
VERSION	B732046.1 GI:16115321
KEYWORDS	EST.
SOURCE	house mouse.
ORGANISM	Mus musculus

## REFERENCE AUTHORS

**AUTHORS**  
Akimura, T., Aikawa, T., Carninci, P., Furuno, M., Hanaoka, T., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Imochi, K., Iishi, Y., Ito, M., Kawai, J., Kojima, Y., Kono, H., Kouda, M., Matsuyama, T., Nakamura, M., Nishi, K., Nomura, K., Nunasaki, R., Okazaki, Y., Sakai, R., Sakai, C., Sakai, K., Sakazume, N., Sasaki, D., Sato, K., Shibata, K., Shinaigawa, A., Shiraki, T., Sogabe, Y., Suzuki, H., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toyu, T., Watabiki, A., Yasunishi, A., Yamamoto, M., and Hayashizaki, Y.

**TITLE**  
The Kiken Encyclopedia of Mouse Full-length cDNAs (Akimura, T., et al. 2001)

JOURNAL  
COMMENT

JOURNAL  
CONTACT  
Unpublished (2001)  
Contact: Yoshihide Hayashizaki  
Laboratory for Genome Exploration Research Group, RIKEN Genomic  
Sciences Center(GSC), Yokohama Institute  
The Institute of Physical and Chemical Research (RIKEN)  
1-7-22 Saito-cho, Tsukuba, Ibaraki, Kanagawa 230-0045, Japan  
Tel: 81-45-503-9222  
Fax: 81-45-503-9216  
Email: genome-res@sc.riken.go.jp,  
URL: <http://genome.gsc.riken.go.jp/>  
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh  
M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.  
Normalization and subtraction of cap-trapper-selected cDNAs to  
prepare full-length cDNA libraries for rapid discovery of new  
genes. *Genome Res.* 10 (10), 1617-1630 (2000)  
Wagl, K., Fujikawa, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E.,  
Watanabe, M., Yoneda, Y., Ishikawa, Y., Ozawa, K., Tanaka, T., Matsura  
S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and  
Hayashizaki, Y.  
RIKEN integrated sequence analysis (RISA) system-384-format  
sequencing pipeline with 384 multicapillary sequencer. *Genome Res.*  
10 (11), 1757-1771 (2000)  
Konno, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P., Sugahara  
Y., and Hayashizaki, Y.  
Computer based methods for the mouse full-length cDNA  
encyclopedia: real-time sequence clustering for construction of a  
nonredundant cDNA library. *Genome Res.* 11 (2), 281-289 (2001)  
Please visit our web site (<http://genome.gsc.riken.go.jp>) for  
further details.  
e mouse tissues.  
Location/Qualifiers  
1.432

```

FEATURES
source
location/Qualifiers
1..432
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="E970020A01"
/clone_1lb="RIKEN full-length enriched, 12 days embryo
whole body"
/tissue_type="whole body"
/dev_stage="12 days embryo"

```

/note="Vector: pSPORT1; Site 1: SalI; Site 2: NotI; This clone is among a rearranged set of 15,247 clones from 11 embryo cDNA libraries (including preimplantation stage embryos from unfertilized egg to blastocyst, embryonic part of E7.5 embryos, extraembryonic part of E7.5 embryos, and E12.5 female mesonephros/gonad) and one newborn ovary cDNA library. Average insert size 1.5 kb. All source libraries are cloned unidirectionally with oligo(dT)-Not primers. References include: (1) Genome-wide expression profiling of mid-gestation placenta and embryo using a 15,000 mouse developmental cDNA microarray, 2000, Proc. Natl. Acad. Sci. U S A, 97: 9127-9132; (2) Large-scale cDNA analysis reveals phased gene expression patterns during preimplantation mouse development, 2000, Development, 127: 1737-1749; (3) Genome-wide mapping of 7.5-day mouse embryos from extraembryonic tissue of expression profiling reveals enrichment in the t-complex and under-representation on the X chromosome, 1998, Hum Mol Genet 7: 1967-1978.

BASE COUNT 100 a 119 c 151 g 62 t

Query Match 18.0%; Score 258; DB 10; Length 432;  
Best Local Similarity 99.3%; Pred. No. 2e-112;  
Matches 408; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1017 GCCGAGATGGCTCTCTACCTCCAGCAGCAAGAGCTGATCAACAGTGGGGCAAT 1076  
DB 20 GCCCAGATGGCTCTCTACCTCCAGCAGCAAGAGCTGATCAACAGTGGGGCAAT 79  
QY 1077 GTCCAGAGTGGAGTGGCAGAGAGCGCGGAGCCAGGCTATGAGACATGAGCAATTC 1136  
DB 80 GTCCAGAGTGGAGTGGCAGAGAGCGCGGAGCCAGGCTATGAGACATGAGCAATTC 139  
QY 1137 TGTGACCGTGGAGCAGCTGGCGCAAAATGTTGCGAACATCGACTCCAGCGGGGGCGCC 1196  
DB 140 TGTGAGCGTGGAGCAGCTGGCGCAAAATGTTGCGAACATCGACTCCAGCGGGGGCGCC 199  
QY 1197 TGGGAGTGAAGAGATGACGACATGCTTTGGATGGGAGAGAGGCAATGGCGGCTGGA 1256  
DB 200 TGGGAGTGAAGAGATGACGACATGCTTTGGATGGGAGAGAGGCAATGGCGGCTGGA 259  
QY 1257 GAGAGAGCGGCTGGAGTGGCAGAGAGTGGCGCTGACAGCCGAGCTGATCCAGA 1316  
DB 260 GAGAGAGCGGCTGGAGTGGCAGAGAGTGGCGCTGACAGCCGAGCTGATCCAGA 319  
QY 1317 GCGCAGACCCGAGAGCGGAGCCAGAGGATGCTGAGATCTCGCAGAGACACCGCGCCA 1376  
DB 320 GCGCAGACCCGAGAGCGGAGCCAGAGGATGCTGAGATCTCGCAGAGACACCGCGCCA 379  
QY 1377 CCAGAGTGGCTTCCGCGCGCGGGAAGTTTCAATTAAGAGTCAAGTG 1427  
DB 380 CCAGAGTGGCTTCCGCGCGCGGGAAGTTTCAATTAAGAGTCAAGTG 430

RESULT 4  
LOCUS AK016235/c 957 bp mRNA linear HTC 19-JAN-2002  
DEFINITION Mus musculus adult male testis cDNA, RIKEN full-length enriched library, clone:493056102:similar to RING-FINGER PROTEIN MORF, full insert sequence.  
ACCESSION AK016235  
VERSION AK016235.1 GI:12854889  
KEYWORDS HTC; CAP trapper;  
SOURCE Mus musculus (strain:C57BL/6J) adult male testis cDNA to mRNA, clone:493056102.  
ORGANISM Mus musculus  
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE 1  
AUTHORS Carninci, P., and Hayashizaki, Y.  
TITLE High-efficiency full-length cDNA cloning

JOURNAL  
MEDLINE  
PUBMED  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
MEDLINE  
PUBMED  
REFERENCE  
AUTHORS  
Shibata, K., Itoh, M., Aizawa, K., Nagata, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kusunagi, T., Tashiro, H., Itoh, M., Sumi, I., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Iizawa, M., Ohara, E., Watanabe, Y., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A., and Hayashizaki, Y.  
RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multiplexed capillary sequencer  
Genome Res. 10 (11), 1757-1771 (2000)  
20530913  
11076861  
4

Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamana, I., Saito, T., Okazaki, Y., Gojibori, T., Bono, H., Kasukawa, T., Saito, R., Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H., Kuehl, P., Lewis, S., Matsuo, Y., Nikaido, I., Pesole, G., Quackenbush, J., Schriml, L. M., Staudt, F., Suzuki, R., Tomita, M., Wagner, J., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N., Carninci, P., de Bonaldo, M. F., Brownstein, M. J., Bult, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hofmann, M., Hume, D. A., Kamiya, M., Lee, N. H., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, J., Mombauts, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seta, T., Shibata, Y., Storch, K. F., Suzuki, H., Toyo-Oka, K., Wang, K. H., Wetz, C., Whitaker, C., Wilming, L., Wyszewski, B., A., Yoshida, K., Hasegawa, Y., Kawai, J., Kohsaki, S., and Hayashizaki, Y.  
Functional annotation of a full-length mouse cDNA collection  
Nature 409 (6821), 685-690 (2001)  
21085660  
11217851

JOURNAL  
MEDLINE  
PUBMED  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
MEDLINE  
PUBMED  
REFERENCE  
AUTHORS  
Shibata, K., Itoh, M., Aizawa, K., Akahira, S., Akimura, T., Aono, H., Araki, A., Arakawa, T., Baldarelli, R., Bono, H., Brownstein, M., Bult, C., Carninci, P., Fukuda, S., Fukunishi, Y., Furuno, M., Hanagaki, T., Hara, A., Hayashizaki, Y., Hori, D., Hirose, K., Hiraoka, T., Hori, F., Hume, D., Imotani, K., Ishii, Y., Itoh, M., Izawa, M., Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Konno, H., Konda, M., Koyama, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Okazaki, Y., Okido, T., Owa, C., Quackenbush, J., Schriml, L., Shibata, K., Shibata, Y., Shinagawa, A., Shiraki, T., Tagabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Tanaka, T., Tejima, Y., Toya, T., Yamamura, T., Yamashita, I., Yasunishi, A., Yoshida, K., Yoshino, M., Muramatsu, M. and Hayashizaki, Y.  
Direct Submission  
Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gs.c.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

COMMENT  
Please visit our web site (<http://genome.gsc.riken.go.jp/>) for further details.





BASE COUNT	43 a	65 c	56 g	31 t
ORIGIN				

[illegible]

Db 91 AGCGGGAGCCAAAGGATGCTGAGGATCTGCGCAGAGACCACCGCGCCACCAAGCTCGGCT 150

OY		1389	TCCCCGCCCGGGAAGTT	140
Db		151	TCCCGCCCCGGGAAGTT	169

RESULT	8
LOCUS	AV006120
DEFINITION	AV006120 392 bp mRNA linear EST 25-AUG-1999 AV006120 Mus musculus C57BL/6J heart Mus musculus CDNA clone 1020006617, mRNA sequence.

ACCESSION	AV006120
VERSION	AV006120.1
	GI:4782970

SOURCE	house mouse.
ORGANISM	Mus musculus

## REFERENCE AUTHORS

TITLE	RIKEN MOUSE ESTS
JOURNAL	Unpublished (1999)
COMMENT	Contact: Chie Owa Genetics Laboratory

RIKEN  
3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan  
Tel.: 81-298-36-9145  
Fax: 81-298-36-9098  
Email: genome-reser@r.riken.go.jp  
Thermostabilization and thermoactivation of thermolabile enzymes by  
trehalose and its application for the synthesis of full length cDNA  
(Proc. Natl. Acad. Sci. U.S.A. 95(2):520-524 (1998))  
Transcriptional sequencing: A method for DNA sequencing using RNA  
polymerase (Proc. Natl. Acad. Sci. U.S.A. 95(7):3455-3460 (1998))  
Please visit our web site (<http://genome.ritc.riken.go.jp>) for  
further details.

```

FEATURES
  source      Location/Qualifiers
1. .392      /organism="Mus musculus"
              /strain="C57BL/6J"

```

```

Query Match      5.2%;   Score 75;   DB 9;   Length 392;
Best Local Similarity 100.0%;   Pred. No. 9.5e-25;
Matches 75;   Conservative 0;   Mismatches 0;   Indels 0;   Gaps 0;

```

1177 GACTTCCAGCCGGCGCGCTGGGATGAAGAGATGACGACATGCTTTCGATGGGAG 1236

Db 138 GACTTCAGCCGGCGCGCTGGGGAATGAACAGGATGACGACATGGCTTTGGATGGGAG 197

QY	1237	GAGGGCAATGCGGGG	125
Db	198	GAGGGCAATGCGGGG	212

Db 198 GAGGCAATGCGGG 212

RESULT 9	LOCUS	DEFINITION
AW918753	AW918753	378 bp mRNA linear EST 25-MAY-2000
	EST530057	Rat gene index, normalized rat, norvegicus, Bento Soares
		<i>Rattus norvegicus</i> cDNA clone K01E57 5' end, mRNA sequence.

ACCESSION	AW918753	GI:8084538
VERSION	AW918753.1	

**SOURCE**  
**ORGANISM**

REFERENCE  
AUTHORS

1 (bases 1 to 378)  
Lee, N.H., Glodok, A., Chandra, I., Mason, T.M., Quackenbush, J.,  
Kerlavage, A.R. and Adams, M.D.

TITLE	Author
Kerlavage, A.R. and Adams, M.J.	
Rat Genome Project: Generation of a Rat EST (RESt) Catalog & Rat	
Gene Index	

JOURNAL  
COMMENT  
Unpublished (1998)  
Contact: lee, NH  
gene index

The Institute for Genomic Research  
9712, Medical Center Drive, Rockville, MD 20850, USA  
Tel: (301)-838-3529  
Fax: (301)-838-0208  
Email: nhlee@tigr.Org  
This clone is available through the ATCC, contact the ATCC  
tel#703-365-2700 for further information  
Seq primer: M13 Reverse

FEATURES	Location/Qualifiers
source	1. .378
	/organism="Rattus norvegicus"

```

/organism="Rattus norvegicus"
/db_xref="taxon:10116"
/clone="RG1857"
/clone_lib="Rat gene index, normalized rat, norvegicus,
Bento Soares"
/tissue_type="mix - brain, ovary, placenta, kidney, lung
liver, embryo, heart, muscle, spleen"
/lab_host="SOLR"
/notes="vector: pbluescript SK(+); Site_1: EcoRI; Site_2:
NotI; Site_3: XbaI"

```

BASE COUNT	87 a	100 c	122 g	69 t
ORIGIN				

Query Match 4.8%; Score 68; DB 10; Length 378;  
 Best Local Similarity 100.0%; Pred. No. 2,1e-21;  
 Matches 68; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 655 GGGCCCAAGAGACTGTGAGGTGGCCCTCTGCCACCATTTACAAAGCCAGAGAGT 714  
 |||||||  
 DB 58 GGGCCCCCAAGAGACTGTGAGGTGGCCCTCTGCCACCATTTACAAAGCCAGAGAGT 117  
 |||||||

QY 715 GAGCTGAG 722  
 |||||||

DB 118 GAGCTGAG 125

RESULT 10  
 BE099923 456 bp mRNA linear EST 13-JUN-2000  
 LOCUS UI-R-BJ1-atl-f-12-0-UI.s1 UI-R-BJ1 Rattus norvegicus cDNA clone  
 DEFINITION  
 BE099923  
 UI-R-BJ1-atl-f-12-0-UI 3', mRNA sequence.  
 ACCESSION BE099923  
 VERSION BE099923.1 GI:8491806  
 KEYWORDS EST.  
 SOURCE Norway rat.  
 ORGANISM Rattus norvegicus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
 Rattus.  
 1 (bases 1 to 456)  
 REFERENCE  
 AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.  
 TITLE Normalization and subtraction: two approaches to facilitate gene  
 JOURNAL discovery  
 MEDLINE Genome Res. 6 (9), 791-806 (1996)  
 COMMENT 97044477  
 Contact: Soares, MB  
 Program for Rat Gene Discovery and Mapping  
 University of Iowa  
 451 Eckstein Medical Research Building Iowa City, IA 52242, USA  
 Tel: 319 335 8250  
 Fax: 319 335 9565  
 Email: msoares@blue.weeg.uiowa.edu  
 The sequence contained an oligo-dT track that was present in the  
 oligonucleotide that was used to prime the synthesis of first  
 strand cDNA and therefore this may represent a bonafide poly A  
 tail. The sequence tag present in the cDNA between the NotI site  
 and the oligo-dT track served to identify it as a clone from the  
 M.B. Soares Lab Clone distribution: clones will be available  
 through Research Genetics (www.resgen.com)  
 Seq primer: M13 Forward  
 POLYA=yes.

FEATURES  
 source  
 1.456  
 Location/Qualifiers  
 /organism="Rattus norvegicus"  
 /strain="Sprague-Dawley"  
 /db\_xref="taxon:10116"  
 /clone="UI-R-BJ1-atl-f-12-0-UI"  
 /clone\_1b="UI-R-BJ1"  
 /lab\_host="DH10B (Life Technologies)"  
 /note="Vector: pRT3D-Pac (Pharmacia) with a modified  
 polylinker: Site\_1: Not I; Site\_2: Eco RI; The UI-R-BJ1  
 library is a subtracted library derived from the following  
 tissues: atrium at 16.5 dpc, ventricle at 16.5 dpc, AV  
 canal at 16.5 dpc, atrium at 15 dpc, ventricle at 15 dpc,  
 AV canal at 15 dpc, ventricle at 13 dpc, and adult heart.  
 For a detailed description of the library from which this  
 clone was derived, please visit our web site at  
 reatest.eng.uiowa.edu. The subtraction has been previously  
 described in (Bonaldo, Lennon and Soares, Genome Research  
 6:791-806, 1996)  
 TAG\_LIB=UI-R-BJ1  
 TAG\_TISSUE=AV canal at 16.5 dpc  
 TAG\_SEQ=GAACC"  
 153 c 114 g 116 t

BASE COUNT 73 a

ORIGIN

Query Match 4.5%; Score 65; DB 10; Length 456;  
 Best Local Similarity 100.0%; Pred. No. 6.3e-20;  
 Matches 65; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 982 AACGTGTGAGTCCGCATCCATGAGAGAGACCCGAGATGCTCTACTCCAG 1041  
 |||||||  
 DB 454 AAGCTGTGAGTCCGCATCCATGAGAGAGACCCGAGATGCTCTACTCCAG 395  
 |||||||

QY 1042 CAGGC 1046  
 |||||||

DB 394 CAGGC 390

RESULT 11  
 A1712700/c 504 bp mRNA linear EST 08-JUN-1999  
 LOCUS UI-R-AF1-aau-e-02-0-UI.s1 UI-R-AF1 Rattus norvegicus cDNA clone  
 DEFINITION  
 A1712700  
 UI-R-AF1-aau-e-02-0-UI 3', mRNA sequence.  
 ACCESSION A1712700  
 VERSION A1712700.1 GI:5016500  
 KEYWORDS EST.  
 SOURCE Norway rat.  
 ORGANISM Rattus norvegicus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
 Rattus.  
 1 (bases 1 to 504)  
 REFERENCE  
 AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.  
 TITLE Normalization and subtraction: two approaches to facilitate gene  
 JOURNAL discovery  
 MEDLINE Genome Res. 6 (9), 791-806 (1996)  
 COMMENT 97044477  
 Contact: Soares, MB  
 Program for Rat Gene Discovery and Mapping  
 University of Iowa  
 451 Eckstein Medical Research Building Iowa City, IA 52242, USA  
 Tel: 319 335 8250  
 Fax: 319 335 9565  
 Email: msoares@blue.weeg.uiowa.edu  
 The sequence contained an oligo-dT track that was present in the  
 oligonucleotide that was used to prime the synthesis of first  
 strand cDNA and therefore this may represent a bonafide poly A  
 tail. The sequence tag present in the cDNA between the NotI site  
 and the oligo-dT track served to verify it as a clone from the  
 M.B. Soares Lab Clone distribution: clones will be available  
 through Research Genetics (www.resgen.com)  
 Seq primer: M13 Forward  
 POLYA=yes.

FEATURES  
 source  
 1.504  
 Location/Qualifiers  
 /organism="Rattus norvegicus"  
 /strain="Sprague-Dawley"  
 /db\_xref="taxon:10116"  
 /clone="UI-R-AF1-aau-e-02-0-UI"  
 /clone\_1b="UI-R-AF1"  
 /dev\_stage="adult"  
 /lab\_host="DH10B (Life Technologies)"  
 /note="Vector: pRT3D-Pac (Pharmacia) with a modified  
 polylinker: Site\_1: Not I; Site\_2: Eco RI; The UI-R-AF1  
 library is a normalized library constructed from 15 dpc  
 rat atrioventricular (AV) canal. The tag is a string of 5  
 nucleotides present between the Not I site and the  
 oligo-dT track. The library was constructed as described  
 by Bonaldo, Lennon and Soares, Genome Research 6: 791-806  
 , 1996. Tissue provided by Jim Lin, Department of Biology,  
 University of Iowa.  
 TAG\_LIB=UI-R-AF1  
 TAG\_TISSUE=AV canal at 15 dpc  
 TAG\_SEQ=CAAGC"  
 166 c 132 g 124 t

BASE COUNT 82 a

[illegible]

DEFINITION	EST189742 Normalized rat heart, Bento Soares Rattus sp. cDNA clone
ACCESSION	R84889.3 end, mRNA sequence.
VERSION	AA800245
KEYWORDS	AA800245.1 GI:2863200
SOURCE	EST.
ORGANISM	Rattus sp.
REFERENCE	Rattus sp.
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
TITLE	1 (bases 1 to 638)
JOURNAL	Lee,N.H., Glodok,A., Chandra,I., Mason,T.M., Quackenbush,J., Kerlavage,A.R. and Adams,M.D.
COMMENT	Rat Genome Project: Generation of a Rat EST (RESr) Catalog & Rat Gene Index Unpublished (1998) Contact: Lee, NH The Institute for Genomic Research 9712, Medical Center Drive, Rockville, MD 20850, USA Tel: (301)-838-3529 Fax: (301)-838-0208 Email: nhlee@tigr.org Seq primer: M13-21. location/qualifiers 1. .638 /organism="Rattus sp." /db_xref="arcc (inhost):2007238" /db_xref="taxon:10118" /clone="R84889" /clone_11b="Normalized rat heart, Bento Soares" /note="Organ: heart; Vector: pT73pac; Site_1: EcoRI; Site_2: NotI"
FEATURES	
source	
BASE COUNT	105 a 221 c 165 g 147 t
ORIGIN	
Query Match	4.5%, Score 65; DB 9; Length 638;
Best Local Similarity	100.0%; Pred. No. 7; Le-20;
Matches	65; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY	982 AAGCGTGTGATCGCCATCGCATCGATGAGAGCGCGAGATGCTCTACCTCCAG 1041
DB	439 AAGCTGTGATCGCCATCGCATCGATGAGAGCGCGAGATGCTCTACCTCCAG 380
OY	1042 CAGCG 1046
DB	379 CAGCG 375
RESULT 14	
AZ731346	AZ731346 449 bp DNA linear GSS 25-JAN-2001
LOCUS	RPCT-24-117M3.TV RPCT-24 Mus musculus genomic clone RPCT-24-117M3, DNA sequence.
ACCESSION	AZ731346
VERSION	AZ731346.1 GI:12493317
KEYWORDS	GSS.
SOURCE	house mouse.
ORGANISM	Mus musculus
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 449)
AUTHORS	Zhao,S., Niernan,W., Malek,J., Shatsman,S., Akınret,B., Levins,M., Tseayay,G., Geer,K.K., Krol,M., Shvartsbeyn,A., Gebregeorgis,E., Russell,I.D., de Jong,P. and Fraser,C.M.
TITLE	Mouse BAC End Sequences from Library RPCT-24
JOURNAL	Unpublished (1999)
COMMENT	Other GSSs: RPCT-24-117M3.TV Contact: Shaying Zhao Department of Eukaryotic Genomics The Institute for Genomic Research 9712 Medical Center Dr., Rockville, MD 20850, USA Tel: 301 838 0200 Fax: 301 838 0208



Email: szhao@ligr.org  
 Clones are derived from the mouse BAC library RPCI-24. For BAC library availability, please contact Pieter de Jong (pje@genemail.chi.org). Clones may be purchased from BACPAC Resources (<http://www.chori.org/bacpac/orderingframe.htm>). BAC end plate: [http://www.tigr.org/tdb/bac\\_ends/mouse/bac\\_end\\_intro.html](http://www.tigr.org/tdb/bac_ends/mouse/bac_end_intro.html)  
 Plate: 117 row: M column: 3  
 Seq primer: 17  
 Class: BAC ends.

# FEATURES

## source

Location/Qualifiers  
 1. 449

/organism="Mus musculus"  
 /strain="C57BL/6J"  
 /db\_xref="taxon:10090"  
 /clone="RPCI-24-117M3"  
 /clone.lib="RPCI-24"  
 /sex="Male"  
 /cell-type="Spleen/Brain"  
 /note="Vector: pTARBAC1; Site\_1: BamHI; Site\_2: BamHI; RPCI-24 Mouse BAC library produced by Pieter de Jong. The library was cloned in the pTARBAC1 cloning vector at the BamHI sites using MboI partially digested male C57BL/6J DNA."

BASE COUNT 110 a 110 c 125 g 104 t

## ORIGIN

Query Match 3.7%; Score 53; DB 17; Length 449;  
 Best Local Similarity 100.0%; Pred. No. 3,5e-14;

Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 659 CCCACAGAGCTGTGAGTGGCCCTGCTGCCACCATTTACAAAGCCGAGAG 711

|||||  
 25 CCCACAGAGCTGTGAGTGGCCCTGCTGCCACCATTTACAAAGCCGAGAG 77

## RESULT 15

### AV361069

LOCUS AV361069 184 bp mRNA linear EST 13-NOV-1999  
 DEFINITION AV361069 RIKEN full-length enriched, adult male eyeball Mus  
 musculus cDNA clone 7530421L22 3', mRNA sequence.

ACCESSION AV361069  
 VERSION AV361069.1 GI:6408357

KEYWORDS EST.  
 SOURCE house mouse.

## ORGANISM

Mus musculus  
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 184)

## REFERENCE

### AUTHORS

Kono, H., Aizawa, K., Akahira, S., Akiyama, J., Carninci, P., Endo, T., Fukuda, S., Fukunishi, Y., Hara, A., Hayatsu, N., Hirozane, T., Hori, F., Ishii, Y., Ishikawa, T., Itoh, M., Izawa, M., Kadota, K., Kagawa, I., Kai, C., Kawai, J., Kikuchi, N., Kojima, Y., Koya, S., Kusakabe, M., Matsuyama, T., Miki, R., Mizuno, Y., Nakamura, M., Oda, H., Okazaki, Y., Owa, C., Ozawa, Y., Saito, H., Sano, M., Sato, K., Shibata, K., Shibata, Y., Shigemoto, Y., Shiraki, T., Sogabe, Y., Sugihara, Y., Suzuki, H., Suzuki, A., Watanabe, S., Yamamura, T., Yasunishi, A., Yokota, T., Yoshiki, A., Yoshino, M., Muramatsu, M., and Hayashizaki, Y.  
 RIKEN Mouse ESTs (Kono, H., et al. 1999)  
 Unpublished (1999)

## JOURNAL

### TITLE

CONTACT: Yoshihide Hayashizaki  
 Laboratory for Genome Exploration Research Group, RIKEN Genomic  
 Sciences Center (GSC), Yokohama Institute  
 The Institute of Physical and Chemical Research (RIKEN)  
 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan  
 Tel: 81-45-503-9222  
 Fax: 81-45-503-9216

Email: genome-res@sc.riken.go.jp,  
 URL: <http://genome.gsc.riken.go.jp/>

Sasaki, N., Izawa, M., Watanabe, M., Okazaki, Y., and Hayashizaki, Y.  
 Matsunaga, S., Carninci, P., Muramatsu, M., Okazaki, Y. and Hayashizaki, Y.

Transcriptional sequencing: A method for DNA sequencing using RNA

polymerase. Proc. Natl. Acad. Sci. U.S.A. 95 (7), 3455-3460 (1998)  
 Itoh, M., Katsunai, T., Akiyama, J., Shibata, K., Izawa, M., Kawai, J., Tomaru, Y., Carninci, P., Shibata, Y., Ozawa, Y., Muramatsu, M., Okazaki, Y. and Hayashizaki, Y.

Automated filtration-based high-throughput plasmid preparation system. Genome Res. 9 (5), 463-470 (1999)  
 Carninci, P. and Hayashizaki, Y.  
 High-efficiency full-length cDNA cloning. Methods Enzymol. 303, 19-44 (1999)  
 Please visit our web site (<http://genome.rtc.riken.go.jp/>) for further details.

## FEATURES

### source

Location/Qualifiers  
 1. 184

/organism="Mus musculus"  
 /strain="C57BL/6J"  
 /db\_xref="taxon:10090"  
 /clone="7530421L22"  
 /clone.lib="RIKEN full-length enriched, adult male eyeball"  
 /sex="male"  
 /tissue-type="eyeball"  
 /dev\_stage="adult"  
 /lab\_host="DH10B"  
 /note="Site\_1: SalI; Site\_2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was prepared with a primer [5' GAGAGAGAGATTCACAGACTCTTTTCTTTTCTTTTCTTTVN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 10.0 and subtraction to Rot = 185.0. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGAGATTCACAGACTCTTTTCTTTTCTTTTCTTTVN 3']. cDNA was cloned into the XhoI and BamHI sites. Vector: a modified plasmid (KS(+)) after bulk excision from Lambda FLC I. Cloning sites, 5' end: SalI; 3' end: BamHI"

BASE COUNT 44 a 61 c 52 g 27 t

## ORIGIN

Query Match 3.6%; Score 52; DB 10; Length 184;  
 Best Local Similarity 100.0%; Pred. No. 7,5e-14;

Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1356 TGGCAGAGAGACACCGCGCCAGCTGCTCCCGCCCGGAGAGTT 1407

|||||  
 Db 107 TGGCAGAGAGACACCGCGCCAGCTGCTCCCGCCCGGAGAGTT 158

## RESULT 16

LOCUS BG375657 440 bp mRNA linear EST 12-MAR-2001  
 DEFINITION BG375657 UI-R-CS0-btd-h-03-0-UI-s1 UI-R-CS0 Rattus norvegicus cDNA clone  
 UI-R-CS0-btd-h-03-0-UI 3', mRNA sequence.

ACCESSION BG375657  
 VERSION BG375657.1 GI:13300129

## KEYWORDS

### EST.

### ORGANISM

Rattus norvegicus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
 Rattus.  
 1 (bases 1 to 440)

REFERENCE Bonaldo, M.F., Lennon, G. and Soares, M.B.  
 Normalization and subtraction: two approaches to facilitate gene  
 discovery  
 Genome Res. 6 (9), 791-806 (1996)

## JOURNAL

### COMMENT

CONTACT: Soares, MB

Program for Rat Gene Discovery and Mapping  
University of Iowa  
451 Eckstein Medical Research Building Iowa City, IA 52242, USA  
Tel: 319 335 8250  
Fax: 319 335 9565  
Email: mscares@blue.weeg.uiowa.edu

The sequence contained an oligo-dT track that was present in the oligonucleotide that was used to prime the synthesis of first strand cDNA and therefore this may represent a bonafide poly A tail. The sequence tag present in the cDNA between the NotI site and the oligo-dT track served to verify it as a clone from the normalized rat heart pool library cDNA library preparation: M.B. Soares Lab Clone Distribution: clones will be available through Research Genetics (www.resgen.com)  
Seq primer: M13 Forward  
POLYA=Yes.

#### FEATURES

##### Source

Location/Qualifiers  
1..440  
/organism="Rattus norvegicus"  
/strain="Sprague-Dawley"  
/db\_xref="taxon:10116"  
/clone="UI-R-CS0-btd-h-03-0-UI"  
/clone\_11b="UI-R-CS0"  
/dev\_stage="ADULT"  
/lab\_host="DH10B (Life Technologies)"  
/note="Vector: pT73D-Pac (Pharmacia) with a modified polylinker. Site.1: Not I; Site.2: Eco RI; The UI-R-CS0 library is a normalized library constructed from the following rat heart tissues: embryonic day 17, embryonic day 19, embryonic day 21, adult day 1, adult day 12, adult day 75, adult day 200. For a detailed description of the library from which this clone was derived, please visit our web site at [ratest.eng.uiowa.edu](http://ratest.eng.uiowa.edu). The subtraction has been previously described in (Bonaldo, Lennon and Soares, Genome Research 6:791-806, 1996)  
TAG\_11b=UI-R-CS0  
TAG\_RISUSE=rat heart pool  
TAG\_SEQ=ATAGATAC"

BASE COUNT 69 a 144 c 113 g 114 t  
ORIGIN

Query Match 3.4%; Score 48; DB 12; Length 440;  
Best Local Similarity 100.0%; Pred. No. 8.7e-12;  
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 999 CATTGAGTCCATGAGAGCGCAGATGCTCTACTCCAGCAGGC 1046  
|||||  
Db 440 CATTGAGTCCATGAGAGCGCAGATGCTCTACTCCAGCAGGC 393

RESULT 17 441 bp mRNA linear EST 12-MAR-2001  
BG379943  
LOCUS UI-R-CS0-bto-b-01-0-UI.s1 UI-R-CS0 Rattus norvegicus cDNA clone  
DEFINITION UI-R-CS0-bto-b-01-0-UI 3', mRNA sequence.  
ACCESSION BG379943  
VERSION BG379943.1 GI:13304415  
KEYWORDS EST.  
SOURCE Norway rat.  
ORGANISM Rattus norvegicus  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.  
1 (bases 1 to 441)  
Bonaldo, M.F., Lennon, G. and Soares, M.B.  
Normalization and subtraction: two approaches to facilitate gene  
discovery  
Genome Res. 6 (9), 791-806 (1996)

REFERENCE 1 (bases 1 to 441)  
Bonaldo, M.F., Lennon, G. and Soares, M.B.  
Normalization and subtraction: two approaches to facilitate gene  
discovery  
Genome Res. 6 (9), 791-806 (1996)  
JOURNAL MEDLINE  
MEDLINE 97044477  
CONTACT: Soares, MB  
PROGRAM for Rat Gene Discovery and Mapping  
UNIVERSITY of Iowa  
451 Eckstein Medical Research Building Iowa City, IA 52242, USA

#### FEATURES

##### Source

Location/Qualifiers  
1..441  
/organism="Rattus norvegicus"  
/strain="Sprague-Dawley"  
/db\_xref="taxon:10116"  
/clone="UI-R-CS0-bto-b-01-0-UI"  
/clone\_11b="UI-R-CS0"  
/dev\_stage="ADULT"  
/lab\_host="DH10B (Life Technologies)"  
/note="Vector: pT73D-Pac (Pharmacia) with a modified polylinker. Site.1: Not I; Site.2: Eco RI; The UI-R-CS0 library is a normalized library constructed from the following rat heart tissues: embryonic day 17, embryonic day 19, embryonic day 21, adult day 1, adult day 12, adult day 75, adult day 200. For a detailed description of the library from which this clone was derived, please visit our web site at [ratest.eng.uiowa.edu](http://ratest.eng.uiowa.edu). The subtraction has been previously described in (Bonaldo, Lennon and Soares, Genome Research 6:791-806, 1996)  
TAG\_SEQ=None found"

BASE COUNT 69 a 148 c 110 g 114 t  
ORIGIN

Query Match 3.3%; Score 47; DB 12; Length 441;  
Best Local Similarity 100.0%; Pred. No. 2.6e-11;  
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1048 AAGGAGCTGATCAGACAGTGGGGCATGTGCGAGGTGGCG 1094  
|||||  
Db 396 AAGGAGCTGATCAGACAGTGGGGCATGTGCGAGGTGGCG 350

RESULT 18 250 bp mRNA linear EST 05-NOV-1999  
AV268013  
LOCUS AV268013 RIKEN full-length enriched, adult male testis (DH10B) Mus  
DEFINITION musculus cDNA clone 4930532616 3', mRNA sequence.  
ACCESSION AV268013  
VERSION AV268013.1 GI:6256050  
KEYWORDS EST.  
SOURCE house mouse.  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 250)

REFERENCE 1 (bases 1 to 250)  
Konno, H., Aizawa, K., Akahira, S., Akiyama, J., Carninci, P., Endo, T., Fukuda, S., Fukunishi, Y., Hara, A., Hayatsu, N., Hirozane, T., Horii, F., Ishii, Y., Ishikawa, T., Itoh, M., Izawa, M., Kadota, K., Kagawa, I., Kai, C., Kawai, J., Kikuchi, N., Kojima, Y., Koya, S., Kusakabe, M., Matsuyama, T., Miki, R., Mizuno, Y., Nakamura, M., Oda, H., Okazaki, Y., Owa, C., Ozawa, Y., Saito, H., Sano, M., Sato, K., Shibata, K., Shibata, Y., Shigemoto, Y., Shiraki, T., Sogabe, Y., Sugahara, Y., Suzuki, H., Suzuki, H., Takahashi, F., Tateo, M., Tomioka, N., Tsunoda, Y., Watahiki, A., Watanabe, S., Yamamura, T., Yasunishi, A., Yokota, T., Yoshiki, A., Yoshino, M., Muramatsu, M. and Hayashizaki, Y.  
RIKEN Mouse ESTs (Konno, H., et al. 1999)  
Unpublished (1999)

TITLE JOURNAL  
CONTACT: Yoshinide Hayashizaki  
LABORATORY for Genome Exploration Research Group, RIKEN Genomic  
Sciences Center (GSC), Yokohama Institute  
The Institute of Physical and Chemical Research (RIKEN)  
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan  
Tel: 81-45-503-9222

Fax: 81-45-503-9216  
 Email: genome-res@sc.riken.go.jp  
 URL: //genome.gsc.riken.go.jp/  
 Sasaki, N., Itawa, M., Watabiki, M., Ozawa, K., Tanaka, T., Yoneda, Y., Matsura, S., Carninci, P., Muramatsu, M., Okazaki, Y. and Hayashizaki, Y.  
 Transcriptional sequencing: A method for DNA sequencing using RNA polymerase. Proc. Natl. Acad. Sci. U.S.A. 95 (7), 3455-3460 (1998)  
 Itoh, M., Katsunai, T., Akiyama, J., Shibata, K., Itawa, M., Kawai, J., Tomaru, Y., Carninci, P., Shibata, Y., Ozawa, Y., Muramatsu, M., Okazaki, Y. and Hayashizaki, Y.  
 Automated filtration-based high-throughput plasmid preparation system. Genome Res. 9 (5), 463-470 (1999)  
 Carninci, P. and Hayashizaki, Y.  
 High-efficiency full-length cDNA cloning. Methods Enzymol. 303, 19-44 (1999)  
 Please visit our web site (<http://genome-rtc.riken.go.jp>) for further details.

## FEATURES

Source  
 Location/Qualifiers  
 1. .250  
 /organism="Mus musculus"  
 /strain="C57BL/6J"  
 /db\_xref="taxon:10090"  
 /clone="4930532616"  
 /clone\_lib="RIKEN full-length enriched, adult male testis (DH10B)"  
 /sex="male"  
 /tissue\_type="testis"  
 /dev\_stage="adult"  
 /lab\_host="DH10B"  
 /note="Site 1: SalI; Site 2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5', GAGGAGAGAGAGATCCAGAGACCTTTTCTTTTCTTNN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. Second strand cDNA was prepared with the primer adapter of sequence [5', GAGGAGAGATTCGACTTATTTATTAATATCCCCCCCCC 3']. cDNA was cloned into the XhoI and BamHI sites. Vector: a modified Bluescript KS(+) after bulk excision from Lambda FLX I. Cloning sites, 5' end: SalI; 3' end: BamHI."  
 BASE COUNT 65 a 66 c 75 g 44 t  
 ORIGIN  
 Query Match 3.2%; Score 46; DB 9; Length 250;  
 Best Local Similarity 100.0%; Pred. No. 6.4e-11;  
 Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1306 TCTGATCCAGAGCGACCGAGCGGAGCGGATGCTGAG 1351  
 Db 127 TCTGATCCAGAGCGACCGAGCGGAGCGGATGCTGAG 172  
 RESULT 19  
 BE095943/c 386 bp mRNA linear EST 12-JUN-2000  
 LOCUS BE095943  
 DEFINITION UT-R-BU0-apt-e-03-0-UI s1 UT-R-BU0 Rattus norvegicus cDNA clone  
 ACCESSION BE095943  
 VERSION BE095943.1  
 KEYWORDS GI:8486874  
 SOURCE EST.  
 ORGANISM Norway rat.  
 Rattus norvegicus  
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 1 (bases 1 to 386)  
 Bonaldo, M.F., Lennon, G. and Soares, M.B.

TITLE  
 JOURNAL  
 MEDLINE  
 COMMENT  
 Normalization and subtraction: two approaches to facilitate gene discovery  
 Genome Res. 6 (9), 791-806 (1996)  
 97044477  
 Contact: Soares, MB  
 Program for Rat Gene Discovery and Mapping  
 University of Iowa  
 451 Eckstein Medical Research Building Iowa City, IA 52242, USA  
 Tel: 319 335 8250  
 Fax: 319 335 9565  
 Email: msoares@blue.weeg.uiowa.edu

The sequence contained an oligo-dT track that was present in the oligonucleotide that was used to prime the synthesis of first strand cDNA and therefore this may represent a bonafide poly A tail. The sequence tag present in the cDNA between the NotI site and the oligo-dT track served to identify it as a clone from the normalized ganglia library cDNA Library Preparation: M.B. Soares Lab Clone distribution: clones will be available through Research Genetics ([www.resgen.com](http://www.resgen.com))  
 Seq primer: M13 Forward  
 POLA=yes

## FEATURES

Source  
 Location/Qualifiers  
 1. .386  
 /organism="Rattus norvegicus"  
 /strain="Sprague-Dawley"  
 /db\_xref="taxon:10116"  
 /clone="UT-R-BU0-apt-e-03-0-UI"  
 /clone\_lib="UT-R-BU0"  
 /dev\_stage="adult"  
 /lab\_host="DH10B (Life Technologies)"  
 /note="Vector: pTZ19-Pac (Pharmacia) with a modified polylinker. Site 1: Not I; Site 2: Eco RI; The UT-R-BU0 library is a subtracted library derived from a mixture of eye and ganglia tissues. For a detailed description of the library from which this clone was derived, please visit our web site at [ratest.eng.uiowa.edu](http://ratest.eng.uiowa.edu). The subtraction has been previously described in (Bonaldo, Lennon and Soares, Genome Research 6:791-806, 1996)  
 TAG LIB-UT-R-BU0  
 TAG-TISSUE=ganglia  
 TAG-SEQ-GCAGAC"  
 BASE COUNT 61 a 131 c 93 g 101 t  
 ORIGIN  
 Query Match 3.2%; Score 46; DB 10; Length 386;  
 Best Local Similarity 100.0%; Pred. No. 7.5e-11;  
 Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1096 GGACGGCGGAGCGAGCGATGAGAGCATGAGCATCTCTGTGA 1141  
 Db 346 GGACGGCGGAGCGAGCGATGAGAGCATGAGCATCTCTGTGA 301

RESULT 20  
 BM087315 467 bp mRNA linear EST 19-NOV-2001  
 LOCUS BM087315  
 DEFINITION 500054 MARC 2B0V Bos taurus cDNA 5', mRNA sequence.  
 ACCESSION BM087315  
 VERSION BM087315.1  
 KEYWORDS GI:16997943  
 SOURCE EST.  
 ORGANISM Bos taurus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovidae; Bovinae; Bos.  
 1 (bases 1 to 467)  
 Smith, T.P.L., Grose, W.M., Freking, B.A., Roberts, A.J., Stone, R.T., Casas, E., Wray, J.E., White, J., Cho, J., Fahrnenkrug, S.C., Bennett, G.L., Heaton, M.P., Laegreid, W.W., Rohrer, G.A., Chitko-Mckown, C.G., Pertea, G., Holt, I., Karamycheva, S., Liang, F., Quackenbush, J. and Keefe, J.W.  
 Sequence evaluation of four pooled-tissue normalized bovine cDNA libraries and construction of a gene index for cattle

JOURNAL Genome Res. 11 (4), 626-630 (2001)  
 MEDLINE 21180013  
 COMMENT

Contact: Smith TPL  
 USDA, ARS, US Meat Animal Research Center  
 PO Box 166, Clay Center, NE 68933-0166, USA  
 Tel: 402 762 4366  
 Fax: 402 762 4390

Email: smith@email.marc.usda.gov  
 Single pass sequencing. Bases called and alt trimmed with phred  
 v0.980904.e. Vector identified by cross\_match with the -minscore 18  
 and -mismatch 12 options.

PCR primers  
 FORWARD: AGGAACACGCTATGACCAT  
 BACKWARD: GTTTCCTCCAGTCACGACG  
 Plate: 137 row: H column: 12  
 Seq primer: ATTAGGTGACACTATAG.

FEATURES  
 source Location/Qualifiers

1..467  
 /organism="Bos taurus"  
 /db\_xref="taxon:9913"  
 /clone\_id="MARC 2B0Y"  
 /tissue\_type="pooled"  
 /lab\_host="DH10B"  
 /note="Vector: pCMV SPORT6; Site\_1: NotI; Site\_2: SalI;  
 library made from pooled tissue from testis, thymus,  
 semitendinosus muscle, longissimus muscle, pancreas,  
 adrenal, and endometrium."  
 BASE COUNT 113 a 112 c 172 g 70 t  
 ORIGIN

Query Match 3.2%; Score 46; DB 13; Length 467;  
 Best Local Similarity 100.0%; Pred. No. 8.1e-11;  
 Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 257 TGGAGAGACAGCTATTGCCCCATCTGCTGGAGAGATGTTCTCCAA 302  
 ||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 236 TGGAGAGACAGCTATTGCCCCATCTGCTGGAGAGATGTTCTCCAA 281

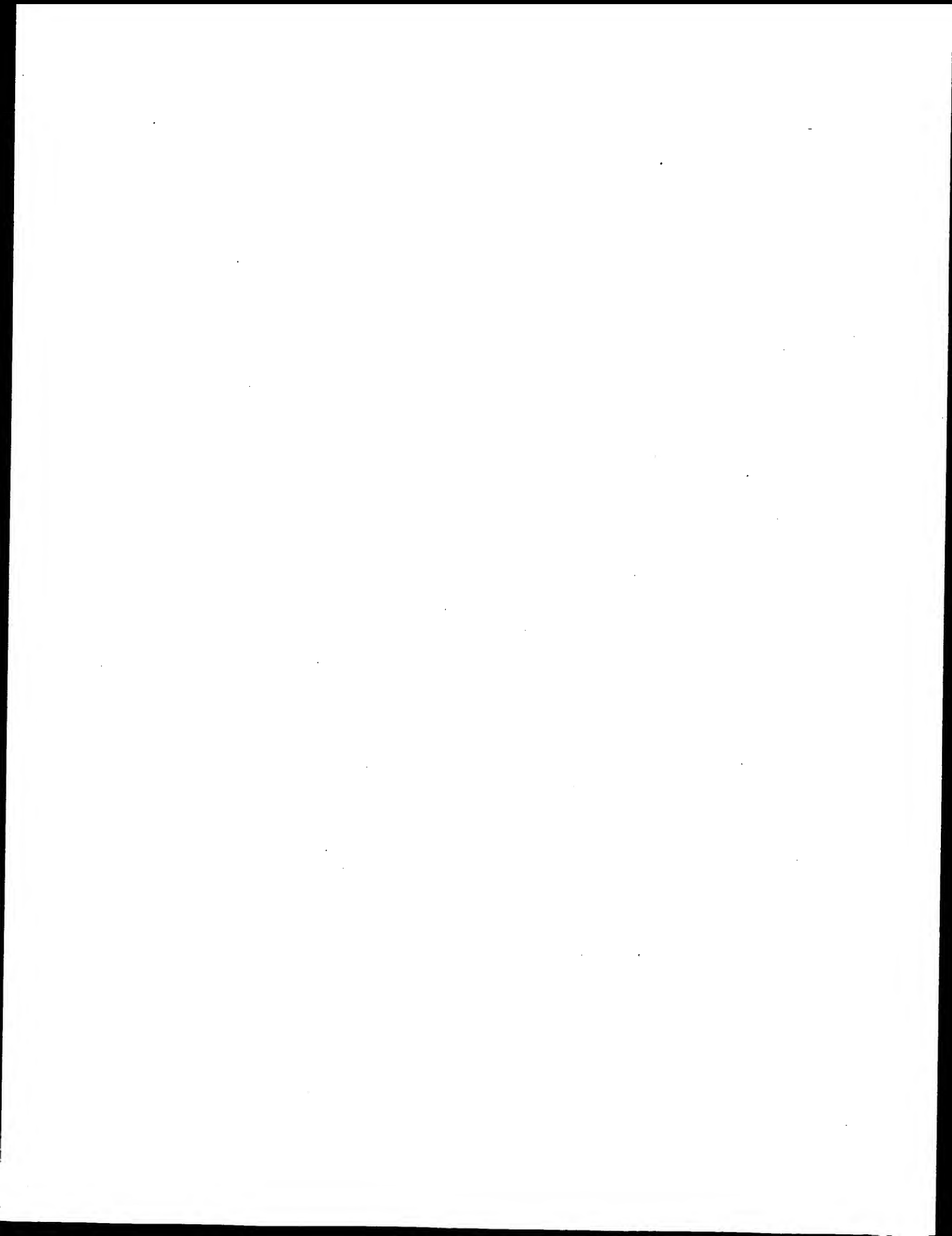
Search completed: December 3, 2002, 12:05:11  
 Job time: 2081 secs



LOCUS AA800245 638 bp mRNA linear EST 30-APR-1998  
 DEFINITION EST189742 Normalized rat heart, Bento Soares Rattus sp. cDNA clone  
 RHEAM89 3' end, mRNA sequence.  
 ACCESSION AA800245  
 VERSION AA800245.1 GI:2863200  
 KEYWORDS EST.  
 SOURCE Rattus sp.  
 ORGANISM Rattus sp.  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
 Rattus.  
 REFERENCE 1 (bases 1 to 638)  
 AUTHORS Lee, N.H., Glodek, A., Chandra, I., Mason, T.M., Quackenbush, J.,  
 Kerlavage, A.R. and Adams, M.D.  
 TITLE Rat Genome Project: Generation of a Rat EST (REST) Catalog & Rat  
 Gene Index  
 JOURNAL Unpublished (1998)  
 COMMENT Contact: Lee, NH  
 The Institute for Genomic Research  
 9712, Medical Center Drive, Rockville, MD 20850, USA  
 Tel: (301)-838-3529  
 Fax: (301)-838-0208  
 Email: nhlee@tigr.org  
 Seq primer: M13-21.  
 FEATURES  
 source Location/Qualifiers  
 1. 638  
 /organism="Rattus sp."  
 /db\_xref="ATCC (inhost):2007238"  
 /db\_xref="taxon:10118"  
 /clone="RHEAM89"  
 /clone\_lib="Normalized rat heart, Bento Soares"  
 /note="Organ: heart; Vector: pT7T3Pac; Site\_1: EcoRI;  
 Site\_2: NotI"  
 BASE COUNT 105 a 221 c 165 g 147 t  
 ORIGIN

Query Match 37.6%; Score 538.2; DB 9; Length 638;  
 Best Local Similarity 91.7%; Pred. No. 2.4e-107;  
 Matches 595; Conservative 0; Mismatches 43; Indels 11; Gaps 2;

Qy 783 GGAGGAGGTGTGCCAGACCATTGAGGACAACAGCCGACAGACAGAAGCAACTGTTAAACCA 842  
 |||  
 Db 638 GGAGGAGGTGTGCCAGACCATTGAGGACAACAGCCGACAGACAGAAACAACTGTTAAACCA 579  
 Qy 843 GAGGTTTCGAGACCCCTGTGCGCGGTTTTGGAGGAGCGCAAGGGCGAACTGCTTCAAGCACT 902  
 ||  
 Db 578 GAAGTTTCGAGACCCCTGTGCGCGGTTTTGGAGGAGCGCAAGGGCGAGTTGCTTCAAGCGCT 519  
 Qy 903 GGCCCGGGAGCAGGAGGAGAAGTTGCAGCGCGTGCAGGGGCTCATCCGCCAGTACGGAGA 962  
 |||  
 Db 518 GGCTCGGGTGCAAGAGGAGAAGCTGCAGCGCGTGCCTAGCCTCATCCGCCAGTACGGAGA 459  
 Qy 963 CCACTTGGAGGGCTCCTCAAAGCTGGTGGAGTCCGCCATCCAGTCCATGGAGGAGCCGCA 1022  
 |||  
 Db 458 CCACTTGGAGGGCTCCTCTAAGCTGGTGGAGTCCGCCATCCAGTCCATGGAGGAGCCGCA 399  
 Qy 1023 GATGGCTCTCTACCTCCAGCAGGCCAAAGGAGCTGATCAACAAGGTCGGGGCAATGTCGAA 1082  
 |||  
 Db 398 GATGGCTCTCTACCTCCAGCAGGCCAAAGGAGCTGATCAACAAGGTCGGGGCAATGTCGAA 339  
 Qy 1083 GGTGGAGCTGGCAGGACGGCCGAGCCAGGCTATGAGAGCATGGAGCAATTCTCTGTGAG 1142  
 |||  
 Db 338 GGTGGAGCTGGCTGGACGGCCGAGCCAGGCTATGAGAGCATGGAGCAATTCTCTGTGAT 279  
 Qy 1143 CGTGGAGCACGTGGCCGAAATGTTGCGAACCATCGACTTCCAGCCGGGCGCGCTGGGGA 1202  
 |||  
 Db 278 CGTGGAGCACGTGGCCGAAATGTTGCGGACCATCGACTTCCAGCCGGGCGCCTCTGGGGA 219  
 Qy 1203 TGAAGAGGATGACGACATGGCTTTGGATGGGGAGGAGGGCAATGCGGGGCTGGAGGAGGA 1262  
 |||  
 Db 218 TGAAGAGGATGACGAGGTGACTTTGACGCGGGGAAGAGGGCAACACAGGGCTGGAGGAGGA 159  
 Qy 1263 GCGGCTGGACGTGCCAGAAGGCTCAGGCCTGCACTGACCCGACTCTGATCCAGAGCGCAC 1322  
 |||  
 Db 158 GCGGCTGGACGGGCCAGAAGGTT-----TGCACTGATCCGACTCTGATCCAGAGCGCAC 105  
 Qy 1323 ACCCGAAGCGGGAGCCAAGGGATGCTGAGGATCTGCGCAGAGACCACCGCGCCACCAAGC 1382  
 ||  
 Db 104 GCC-----CGGGAGCCAAGGGATGCTGAGGATCTGCGCAGAGACCATCGCGCTACCCATC 50  
 Qy 1383 TCGGCTTCCCGCCCCCGGAAGGTTCTCAATAAAGGACTCAAGTGTCCC 1431  
 |||  
 Db 49 TCGGCCTCCACCCCGAGGATGGTTCTCAATAAAGAACTCATGTGTCCC 1





LOCUS AA800245 638 bp mRNA linear EST 30-APR-1998  
 DEFINITION EST189742 Normalized rat heart, Bento Soares Rattus sp. cDNA clone  
 RHEAM89 3' end, mRNA sequence.  
 ACCESSION AA800245  
 VERSION AA800245.1 GI:2863200  
 KEYWORDS EST.  
 SOURCE Rattus sp.  
 ORGANISM Rattus sp.  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
 Rattus.  
 REFERENCE 1 (bases 1 to 638)  
 AUTHORS Lee, N.H., Glodek, A., Chandra, I., Mason, T.M., Quackenbush, J.,  
 Kerlavage, A.R. and Adams, M.D.  
 TITLE Rat Genome Project: Generation of a Rat EST (REST) Catalog & Rat  
 Gene Index  
 JOURNAL Unpublished (1998)  
 COMMENT Contact: Lee, NH  
 The Institute for Genomic Research  
 9712, Medical Center Drive, Rockville, MD 20850, USA  
 Tel: (301)-838-3529  
 Fax: (301)-838-0208  
 Email: nhlee@tigr.org  
 Seq primer: M13-21.  
 FEATURES  
 source Location/Qualifiers  
 1..638  
 /organism="Rattus sp."  
 /db\_xref="ATCC (inhost):2007238"  
 /db\_xref="taxon:10118"  
 /clone="RHEAM89"  
 /clone\_lib="Normalized rat heart, Bento Soares"  
 /note="Organ: heart; Vector: pT7T3Pac; Site\_1: EcoRI;  
 Site\_2: NotI"  
 BASE COUNT 105 a 221 c 165 g 147 t  
 ORIGIN

Alignment Scores:  
 Pred. No.: 5.93e-75 Length: 638  
 Score: 778.00 Matches: 157  
 Percent Similarity: 94.15% Conservative: 4  
 Best Local Similarity: 91.81% Mismatches: 8  
 Query Match: 40.69% Indels: 2  
 DB: 9 Gaps: 1

US-09-908-988B-2 (1-366) x AA800245 (1-638)

Qy 196 GluGluValCysGlnThrIleGluAspAsnSerArgArgGlnLysGlnLeuLeuAsnGln 215  
 Db 637 GAGGAGGTGTGCCAGACCATTGAGGACAACAGCGGCAGACAGAAACAAGTGTAAACCAG 578  
 Qy 216 ArgPheGluThrLeuCysAlaValLeuGluGluArgLysGlyGluLeuLeuGlnAlaLeu 235  
 Db 577 AAGTTCGAGACCCTGTGCGCGGTTTTGGAGGAGCGCAAGGGCGAGTTGCTTCAAGCGCTG 518  
 Qy 236 AlaArgGluGlnGluGluLysLeuGlnArgValArgGlyLeuIleArgGlnTyrGlyAsp 255  
 Db 517 GCTCGGGTGCAAGAGGAGAAGCTGCAGCGCGTGCCTAGCCTCATCCGCCAGTACGGAGAC 458  
 Qy 256 HisLeuGluGlySerSerLysLeuValGluSerAlaIleGlnSerMetGluGluProGln 275  
 Db 457 CACTTGGAGGCCTCCTCTAAGCTGGTGGAGTCCGCCATCCAGTCCATGGAGGAGCCGCAG 398  
 Qy 276 MetAlaLeuTyrLeuGlnGlnAlaLysGluLeuIleAsnLysValGlyAlaMetSerLys 295  
 Db 397 ATGGCTCTCTACCTCCAGCAGGCCAAGGAGCTGATCAACAAGGTCGGGGCAATGTGGAAG 338  
 Qy 296 ValGluLeuAlaGlyArgProGluProGlyTyrGluSerMetGluGlnPheSerValSer 315  
 Db 337 GTGGAGCTGGCTGGACGGCCGAGCCAGGCTATGAGAGCATGGAGCAATCTCTGTGATC 278  
 Qy 316 ValGluHisValAlaGluMetLeuArgThrIleAspPheGlnProGlyAlaAlaGlyAsp 335  
 Db 277 GTGGAGCACGTGGCGGAAATGTTGCGGACCATCGACTTCCAGCCGGGCGCCTCTGGGGAT 218  
 Qy 336 GluGluAspAspAspMetAlaLeuAspGlyGluGluGlyAsnAlaGlyLeuGluGluGlu 355  
 Db 217 GAAGAGGATGACGAGGTGACTTTGGACGGGGAAGAGGGCAACACAGGGCTGGAGGAGGAG 158  
 Qy 356 ArgLeuAspValProGluGlySerGlyLeuHis 366  
 Db 157 CGGCTGGACGGGCCAGAA-----GGTTTGAC 131



GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - nucleic search, using frame-plus-p2n model

Run on: December 3, 2002, 13:08:10 ; Search time 2056 Seconds  
(without alignments)  
2883.049 Million cell updates/sec

Title: US-09-908-988b-2  
Sequence: 1 MNFTVGFKPLGDHNNMDN.....EGNAGLEERLDVPEGSLH 366

Scoring table:  
BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Zgapop 6.0 , Zgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 16154066 seqs, 8097743376 residues  
Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+2n,model -DEV=xlh  
-O=/cgn2.1/USPRO-SP001/US09908988/runat\_26112002\_112155\_11286/app\_query.fasta\_1.519  
-DB=EST -QFWT=fastlap -SUFFIX=first -MINMATCH=0.1 -LOOPEXT=0 -LIST=45  
-UNITS=bits -START=1 -END=1 -MATRIX=blomsun62 -TRANS=human40.cdi -ALIGN=20 -MODE=LOCAL  
-DOALIGN=200 -THR.SCORERE=500 -MINLEN=0 -MAXLEN=2000000000  
-OUTFMT=ptc -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09908988@cgn2.1.753@runat\_26112002\_112155\_11286 -NCP=6 -ICPU=3  
-NO\_XLPXY -NO\_XMAP -LANG=EUERY -NEG.SCORERE=0 -WAIT -LONGLOG -DEV\_TIME=OUT=120  
-WARN.TIMEOUT=30 -THREHOLD=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estnu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_hic:\*  
9: gb\_estl:\*  
10: gb\_est2:\*  
11: gb\_hic:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: gb\_gss:\*  
18: em\_gss\_hum:\*  
19: em\_gss\_iny:\*  
20: em\_gss\_pln:\*  
21: em\_gss\_vtc:\*  
22: em\_gss\_fun:\*  
23: em\_gss\_mam:\*  
24: em\_gss\_mus:\*  
25: em\_gss\_other:\*  
26: em\_gss\_pro:\*  
27: em\_gss\_rtd:\*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Match	Query Length	DB ID	Description
1	970	50.7	595	13	Bj071661
2	900	47.1	555	12	BF855618
3	885	46.3	912	13	Bj758535
4	863	45.1	516	13	BM030114
5	850	44.5	512	12	BF855619
6	849.5	44.4	868	12	BF855619
7	825.5	43.2	793	13	Bj743358
8	824	43.1	1105	12	Bj743358
9	813	42.5	694	10	BF852026
10	786	41.1	615	13	Bj033623
11	779.5	40.8	815	9	AL668403
12	778	40.7	638	9	AA800245
13	756	39.5	628	13	Bj059953
14	733.5	38.4	1161	12	Bj163322
15	726	38.0	608	12	BF284900
16	702	36.7	495	13	BM488618
17	697	36.5	1065	17	CNS04GLR
18	695.5	36.4	1092	13	Bj114696
19	693	36.2	776	13	Bj114212
20	687	35.9	552	13	Bj072360
21	684.5	35.8	790	12	BF674823
22	683.5	35.7	582	13	Bj070440
23	683	35.7	658	14	BM726338
24	681	35.6	983	12	BF205857
25	670	35.0	579	13	Bj074770
26	668.5	35.0	1081	12	BF309583
27	629	32.9	373	10	AA483766
28	629	32.9	672	10	BB662321
29	627	32.8	378	10	AA918753
30	624	32.6	573	9	AA443443
31	617	32.3	721	12	Bj765856
32	614	32.1	600	13	Bj891964
33	608.5	31.8	842	17	CNS02TVC
34	605	31.6	610	17	Bj500196
35	599.5	31.4	610	17	DR161818
36	596.5	31.2	682	12	Bj769996
37	595	31.1	464	9	Aj644642
38	593	31.0	485	11	AK019655
39	586	30.6	795	12	BE894237
40	570	29.8	637	14	BM890066
41	557.5	29.2	688	9	Aj442216
42	547	28.6	504	9	Aj712700
43	542.5	28.4	531	9	Aj877673
44	536.5	28.1	508	12	BF724411
45	535.5	28.0	538	9	Aj958735

## ALIGNMENTS

RESULT 1  
LOCUS Bj071661 595 bp mRNA linear EST 11-DEC-2001  
DEFINITION laevis cDNA clone Xl095b10 5', mRNA sequence.  
ACCESSION Bj071661  
VERSION Bj071661.1 GI:17501850  
KEYWORDS EST.  
SOURCE African clawed frog.  
ORGANISM Xenopus laevis  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Mesobatrachia; Pipridae; Pipidae; Xenopodinae; Xenopus.  
REFERENCE 1 (bases 1 to 595)  
Kitayama,A., Terasaka,C., Mochii,M., Ueno,N., Shin-I,T. and Kohara



Db 494 TGCCGCAATGTCGACAGAGCTTTCAGGCTCGATTCCTATGAGCAGTCCGGGCG 435  
 Oy 67 SerThrThrValSerSerGlyGlyIarGpheaArgCysProSerCysArgHisGluVal 86  
 Db 434 TCCACCACTGCTCTTCAGAGAGCCGTTTCCCTCCATCGTCAGAGGATGAGGTTGTC 375  
 Oy 87 LeuAspArgHisGlyValTyrGlyLeuGlnArgAsnLeuValGluAsnIleIleAsp 106  
 Db 374 CTGACAGACAGACGCTGTCTACGCGCTCGACGCAACCTGCTAGTGGAGACATTTATCGAC 315  
 Oy 107 IleTyrGlyGlnGluSerSerArgProLeuHisAlaLysAlaGluGlnHisLeuMetCys 126  
 Db 314 ATTTACCAAGCAGAGATCATCCAGCGCTGCACACTCCAGAGCTGAGCAGCAGCTCATGTGC 255  
 Oy 127 GluGluHisGluAspGluLysIleAsnIleTyrCysLeuSerCysGluValProThrCys 146  
 Db 254 GAGGAGCATGAAGAGAGATCATATTACTGCTCGAGTCGAGGTCGACCTGC 195  
 Oy 147 SerLeuCysLysValPheGlyAlaHisLysAspCysGluValAlaProLeuProThrIle 166  
 Db 194 TCTCTCTGCAAGCTCTTCGCTGCCACCAAGACTGTGAGTGGCCCACTGCCACCAT 135  
 Oy 167 TyrLysArgGlnLysSerGluLeuSerAspGlyIleAlaMetLeuValAlaGlyAsnAsp 186  
 Db 134 TACACAGCCCAAGAGATGAGCTCAGCATGCGCATGCTGTGTGCGAGGCAATGAC 75  
 Oy 187 ArgValGlnAlaValIleThrGlnMetGluGlu-ValCysGlnThr 201  
 Db 74 CCGGTGCAAGCAGTATCATCAGATGAGAGCGGTGTGCGACACT 29  
 RESULT 3 912 bp mRNA linear EST 25-SEP-2001  
 LOCUS B1758535 603022861F1 NIH\_MGC\_114 Homo sapiens CDNA clone IMAGE:5193326 5',  
 DEFINITION mRNA sequence.  
 ACCESSION B1758535  
 VERSION B1758535.1 GI:15750113  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
 NIH-MGC http://mgi.nci.nih.gov/  
 1 (bases 1 to 912)  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished (1999)  
 Contact: Robert Strausberg, Ph.D.  
 Email: cgapbds-remail.nih.gov  
 Tissue Procurement: Life Technologies, Inc.  
 CDNA Library Preparation: Life Technologies, Inc.  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov  
 Plate: LLML1483 row: j column: 15  
 High quality sequence start: 25  
 High quality sequence stop: 890.  
 Location/Qualifiers  
 1. 912  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5193326"  
 /clone\_1bp="NIH\_MGC\_114"  
 /lab\_host="DH10B"  
 /note="Organ: brain; Vector: pCMV-SPORT6; Site\_1: NotI;  
 Site\_2: EcoRV (destroyed); RNA source anonymous pool of 6  
 male brains, age range 23-27 yo. Library is oligo-dT  
 primed and directionally cloned (EcoRV site is destroyed  
 upon cloning). Average insert size 1.5 kb, insert size  
 range 1-3 kb. Library is normalized and enriched for  
 full-length clones and was constructed by C. Gruber  
 (Invitrogen). Research Genetics tracking code 019. Note:

BASE COUNT 236 a 232 c 266 g 178 t  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 1,99e-86 Length: 912  
 Score: 885.00 Matches: 166  
 Percent Similarity: 84.15% Conservative: 41  
 Best local Similarity: 67.48% Mismatches: 34  
 Query Match: 46.29% Indels: 5  
 DB: 13 Gaps: 3  
 US-09-908-988b-2 (1-366) x B1758535 (1-912)  
 Oy 10 LeuLeuGlyAspAlaHisAsnMetAspAsnLeuGluLysGlnLeuIleCysProIleCys 29  
 Db 167 CTGATCCAGGATGGGAATCCATGTGGAATCTTGGAACACAGCTGATCTGCCCTATCTCG 226  
 Oy 30 LeuGluMetPheSerLysProValIleLeuProCysGlnHisAsnLeuCysArgLys 49  
 Db 227 CTGAGATGTTTACCAAGCAGCTGTCATCTTGCCGTGCCAGCACAACCTGTCCGGAG 286  
 Oy 50 CysAlaAsnAspValPheGlnAlaSerAsnProLeuTrpGlnSerArgGlySerThrThr 69  
 Db 287 TGTGCCAATGACATCTTCCAGCTGCATAATCCCTACTGACACCGCGGACCTCAGTG 346  
 Oy 70 ValSerSerGlyIarGpheaArgCysProSerCysArgHisGluValIleAspArg 89  
 Db 347 TCCATGTCTGAGGCGCGTTTCGCTGCCACACTGCCGACAGAGTGATCATGATCGT 406  
 Oy 90 HisGlyValTyrGlyLeuGlnArgAsnLeuValGluAsnIleIleAspIleTyrLys 109  
 Db 407 CACGAGGTGTACGGCTCGACAGGAACCTGCTGTGAGAACTCATCGACATCTACAA 466  
 Oy 110 GlnGlu---SerSerArgProLeuHisAlaLysAlaGluGlnHisLeuMetCysGluGlu 128  
 Db 467 CAGGAGCTCTCAGATCGCCCGCTG-----CACAAGGCGAGTCAACCCATGTGAAGAG 520  
 Oy 129 HisGluAspGluLysIleAsnIleTyrCysLeuSerCysGluValProThrCysSerLeu 148  
 Db 521 CACGAAGATGAGAAATCAACATCTACTGCTCAGCTGATGAGTCCACCTGCTCCATG 580  
 Oy 149 CysLysValPheGlyAlaHisLysAspCysGluValAlaProLeuProThrIleTyrLys 168  
 Db 581 TCCAAAGTGTGTGGATCCACAAGGCTCGAGGTGGCCCATTTGACAGAGTGTTCAG 640  
 Oy 169 ArgGlnLysSerGluLeuSerAspGlyIleAlaMetLeuValAlaGlyAsnAspArgVal 188  
 Db 641 GCACAAAAGACTAAGCTGAATATCTGATCTCCATGCTGGCGGGGAATGACGCTGTG 700  
 Oy 189 GlnAlaValIleThrGlnMetGluGluValCysGlnThrIleGluAsp-AsnSerArgAr 208  
 Db 701 CAGACCATCATCTCAGCTGAGGAGATTCCCTGATGACCAAGAGAAACACTCCACA 760  
 Oy 208 GcInLysGlnLeuLeuAsnGlnArgPheGluThrLeuCysAlaValLeuGluGluArgLys 228  
 Db 761 GGTAAAGGAAGAGCTGACGCAAGATTGACCTTATGCTACCTGAGTGAAGAA 820  
 Oy 228 scGlyLeuLeuGlnAlaLeuAlaArgGluGlnGluLysLeuGlnArgValArgGly 248  
 Db 821 AAGTGAGTTCCTCCAGAGGATCACCGCAGAGCGGAGAAAAGCTTAGTTCAT---CGAGG 877  
 Oy 248 yLeuIleArgGlnTyr 253  
 Db 878 CTTATCTCAGCAGATAC 893  
 RESULT 4 516 bp mRNA linear EST 05-NOV-2001  
 LOCUS BM030114 488634 MARC 2BOV Bos taurus CDNA 5', mRNA sequence.  
 DEFINITION BM030114  
 ACCESSION BM030114.1 GI:16743684  
 VERSION EST.  
 KEYWORDS

SOURCE	CCW.
ORGANISM	Bos taurus
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidea; Bovidae; Bovinae; Bos. 1 (bases 1 to 516)
AUTHORS	Smith,T.P.L., Grose,W.M., Freking,B.A., Roberts,A.J., Stone,R.T., Casas,E., Wray,J.E., White,J., Cho,J., Fahrnkung,S.C., Bennett,G.L., Heaton,M.P., Laegreid,W.M., Rohrer,G.A., Chitko-McKown,C.G., Pettes,G., Holt,I., Karamychena,S., Liang,F., Quakenbush,J. and Keefe,J.W.
TITLE	Sequence evaluation of four pooled-tissue normalized bovine cDNA libraries and construction of a gene index for cattle
JOURNAL MEDLINE	Genome Res. 11 (4), 626-630 (2001)
COMMENT	21180013
	Contact: Smith TPL

BASE COUNT	119 a	137 c	171 g	89 t
ORIGIN				

Alignment Scores:	
Pred. No.:	1.95e-84
Score:	863.00
Percent Similarity:	98.84%
Best Local Similarity:	97.67%
Query Match:	45.14%
DB:	13
	Gaps: 0
	Indels: 0
	Mismatches: 2
	Conservative: 2
	Matches: 168
	Length: 516

US-09-908-988B-2 (1-366) x BM030114 (1-516)

OY	111	GlnSerSerArgProLeuHisAlaValSalGluGlnIleHisMetCysGluGluHisGlu	130
Db	1	GAGTCTCCGGGCACATGTCCAAGCGTAGAGCACCTCATGTGTAGAGACATGA	60
OY	131	AspGluLysIleAsnIleTyrCysLeuSerCysGluValDroThrcySerLeucyls	150
Db	61	GATGAGAAATCATATTACTGTCTTCAGTTGGCAATGCCACCTGTCTCTGTCAAG	120
OY	151	ValPheGlyAlaHisLysAspCysGluValAlaProLeuProThrIleTyrLysArgIn	170
Db	121	GCTTTTGTCGCCACAAGGACTGTGAGGTGGCCCCACTGCCACATTTCAAACGCCAG	180
OY	171	LysSerCLeuLeuSerAspGlyIleLeuMetLeuValAlaGlyAsnAspArgValGlnAla	190
Db	181	AAGAAGTAAGTCAAGCATGTGCATTCGGGATCTGTGTGGCCGGCATATGCCCTGTGCAAGCA	240
OY	191	ValIleThrGlnMetGluGluValCysGlnThrIleGluLysPasnSerArgArgGlnIys	210
Db	241	GTCATCAACACAGATGGAGGAGTGCTGCCAAGCATGTAGAGAAAACGCCGAGCGCAAGC	300

QY	211	GlnLeuLeuAaNGlnInArGpHglTThrLeuGysAlaValLeuGlnGlnArgGlySGlyGln	230
Db	301	CAGTGTGCGAACCAGAGAGTTTGAGGGCCGTGTCCACTACTCGAGAGCGGAAGGGTAC	360
QY	231	LeuLeuGlnAlaLeuAlaValArgGlnGlnGlnGlnGlnGlnArgValAlaArgGlyLeuIle	250
Db	361	CTGCGCGACGGCGTGGCCCGGAGCAGAGAGAGAAAGCTGCAGCGAGTCAAGGGGCTCATC	420
QY	251	ArgGlnTrpGlyAspHisLeuGlnGlySerSerLeuGlnValGlnSerAlaIleGlnSer	270
Db	421	CGCCAGTACGGAGACCAACTGTGAGGCTCCCTCTAAAGCTGTGGAGAGTCTGCCATCACTGCC	480
QY	271	MetGlnGlnProGlnMetAlaLeuTyrLeuGlnGln	282
Db	481	ATGAGGAGCCGACGATGTGGGCTCTACTCTCAGACAG	516

RESULT	5
BE855619/c	
LOCUS	BF855619 512 bp mRNA linear EST 16-JAN-2001
DEFINITION	RCC-FN0204-161100-031-e06 FN0204 Homo sapiens CDNA, mRNA sequence.
ACCESSION	BF855619
VERSION	BF855619.1 GI:12243363
KEYWORDS	EST.
SOURCE	human.
ORGANISM	Homo sapiens

REFERENCE	AUTHORS	TITLE
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 512)	Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Britones, M. R., Negai, M. A., da Silva, W. Jr., Zago, M. A., Bordini, S., Costa, F. F., Goldman, G. H., Carvalho, A. F., Matsukuma, A., Bata, G. S., Simpson, D. H., Brunstein, A. B., deOliveira, P. S., Bucher, P., Jongeneel, C. V., O'Hare, M. J., Soares, F., Brentani, R. R., Reis, L. F., de Souza, S. J. and Simpson, A. J.	Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
JOURNAL MEDLINE	Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)	20202463
COMMENT	Contact: Simpson A.J.G.	

```

/Note=Organ: prostate;normal; Vector: puc18; Site:1: Sma1
; Site:2: Sma1; A mini-library was made by cloning
products derived from ORESTES PCR (U.S. letters Patent
application No. 196,716 - Ludwig Institute for Cancer
Research) profiles into the pUC 18 vector. Reverse
transcription of tissue mRNA and cDNA amplification were
performed under low stringency conditions."
104 a 138 c 157 g 113 t

```

Alignment Scores:	
Pred. No.:	5.16e-83
Score:	850.00
Percent Similarity:	96.51%
Length:	512
Matches:	164
Conservative:	2

Best Local Similarity: 95.35% Mismatches: 3  
 Query Match: 44.46% Indels: 3  
 DB: 12 Gaps: 0

US-09-908-988b-2 (1-366) x BF855619 (1-512)

QY 27 ProileCysLeuGluMetPheSerIysProValIleLeuProCysGlnHisAsnLeu 46  
 |||||  
 DB 511 CCCATCTGCTGGAGATGTTCTCAACACAGTGTGATCTGCGCTGCCAACACAACTG 452  
 QY 47 CysArgIysCysAlaAsnAspValPheGlnAlaSerAsnProLeuTrpGlnSerArgIy 66  
 |||||  
 DB 451 TGCCCAATGTGGCAACACAGCTCTTCAGGCTCGAATCTCTATGGCAGTCCGGGGC 392  
 QY 67 SerThrIleValSerSerGlyIArgPheArgCysProSerCysArgHisGluVal 86  
 |||||  
 DB 391 TCCACACAGTGTCTTCCAGAGAGCCCTTCCGCTGCCATCGTGCAGAGCATGAGTGTGC 332  
 QY 87 LeuAspArgHisGlyValIArgIleGlnArgAsnLeuLeuValGluAsnIleIleAsp 106  
 |||||  
 DB 331 CTGACACAGACAGCTGTCTACGGCTGCAGACCAACCTCTAGTGGAGAACATTCGAC 272  
 QY 107 IleIArgIleGlnIleSerSerArgProLeuHisAlaIleGluGlnIleMetCys 126  
 |||||  
 DB 271 ATTACCAACAGAGATCATCCAGGCGCTGCTCCAGGCTGAGCAGCACTCATGTGC 212  
 QY 127 GlnGluHisGluAspGluIleAsnIleIArgCysLeuSerCysGluValProThrCys 146  
 |||||  
 DB 211 GAGGACATGAGAGAGAGATCATATTACTGCTGAGCTGCAGAGGTGCCACCTGC 152  
 QY 147 SerLeuCysIleValPheGlnAlaHisIleAspCysGluValAlaProLeuProThrIle 166  
 |||||  
 DB 151 TCTCTCTGCAAGGCTCTCGGTGCCACAGAGACTGTGAGGTGGCCCACTGCCACCAT 92  
 QY 167 TyrIleArg- -GlnIysSerGluLeuSerAspGlyIleAlaMetLeuValAlaGlyAsn 186  
 |||||  
 DB 91 TACTAACGCGCCCGAAGAGAGAGCTCAGCATGTGAGGTGAGGTGAGGTGAGGTGAG 32  
 QY 186 sPArgValGlnAla-ValIleThrGlnMet 195  
 |||||  
 DB 31 ACCGCGTGCAGAGCATGTGATCACACGCGATG 2  
 RESULT 6 868 bp mRNA linear EST 15-MAY-2001  
 LOCUS BG764060 602737069F1 NIH\_MGC\_49 Homo sapiens cDNA clone IMAGE:4862321 5',  
 DEFINITION mRNA sequence.  
 ACCESSION BG764060  
 VERSION BG764060.1 GI:14074713  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 868)  
 NIH-MGC http://mhc.nci.nih.gov/  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished (1999)  
 Contact: Robert Strausberg, Ph.D.  
 Email: cgaabs-r@mail.nih.gov  
 Tissue Procurement: ATCC/DCMP/DRP  
 cDNA Library Preparation: Ling Hong/Rubin Laboratory  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LNL at:  
 http://image.llnl.gov  
 plate: LICM1722 row: j column: 18  
 high quality sequence stop: 863.  
 Location/Qualifiers  
 1. 868  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"

/clone="IMAGE:4862321"  
 /clone\_id="NIH\_MGC\_49"  
 /tissue\_type="melanotic melanoma, high MDR (cell line)"  
 /lab\_host="DH10B (phage-resistant)"  
 /note="Organ: skin; Vector: pOTB7; Site:1: XhoI; Site:2:  
 EcoRI; cDNA made by oligo-dT priming. Directionally cloned  
 into EcoRI/XhoI sites using the following 5' adaptor:  
 GGCAAGAGG(c). Size selected >500bp for average insert size  
 1.8kb. Library constructed by Ling Hong in the Laboratory  
 of Gerald M. Rubin (University of California, Berkeley)  
 using ZAP-cDNA synthesis kit (Stratagene) and Superscript  
 II RT (Life Technologies). Note: this is a NIH\_MGC  
 Library."
 

BASE COUNT	224 a	222 c	252 g	170 t
ORIGIN				
Alignment Scores:				
Pred. No.:	1,38e-82	Length:	868	
Score:	849.50	Matches:	163	
Percent Similarity:	82.59%	Conservative:	41	
Best Local Similarity:	65.99%	Mismatches:	37	
Query Match:	44.43%	Indels:	6	
DB:	12	Gaps:	2	

US-09-908-988b-2 (1-366) x BG764060 (1-868)

QY 10 LeuLeuGlyAspAlaHisAsnMetAspAsnLeuGluGlnLeuIleCysProIleCys 29  
 |||||  
 DB 138 CTGATCCAGAGATGGGAAATCCCATGGAGACTTGGAGAGAGCTGATCTGCTATCTGC 197  
 QY 30 LeuGluMetPheSerIysProValIleLeuProCysGlnHisAsnLeuCysArgIys 49  
 |||||  
 DB 198 CTGGAGATGTTTACCAAGCAGAGCTGATCTGCTGCGCAGACAACTGTGCCGAGAG 257  
 QY 50 CysAlaAsnAspValPheGlnAlaSerAsnProLeuTrpGlnSerArgIysThr 69  
 |||||  
 DB 258 TCTGCCAATGACATCTTCCAGGCTGCAATCCTACTGAGCAGACCGGGCAGCTGAGT 317  
 QY 70 ValSerSerGlyIleValPheArgCysProSerCysArgHisGluValIleAspArg 89  
 |||||  
 DB 318 TCCAGTCTGTGAGGCGGTTTCCGCTGCCACCTGCGCCAGAGAGTATCATGATCGT 377  
 QY 90 HisGlyValTyrGlyLeuGlnArgAsnLeuValGluAsnIleIleAspIleTyrIys 109  
 |||||  
 DB 378 CACGAGATGTAGGCGCTGCAAGAGAACTGCTGTGAGACATCATCGATCTACAAA 437  
 QY 110 GlnGlu- -SerSerArgProLeuHisAlaIleGlnIleMetCysGluGlu 128  
 |||||  
 DB 438 CAGGATGCTCCAGTCCGCGCGCTG- -CAGAAGGCGAGTCAACCCAGTGTCAAGAGAG 491  
 QY 129 HisGluAspGluIleAsnIleTyrCysLeuSerCysGluValAlaProLeuProThrIleTyr 148  
 |||||  
 DB 492 CACGAAGATGAGAAATCAAGATCTACTGCTTCACCTGTGAGGTGCGCACCTGCTCAT 551  
 QY 148 ucYsIysValPheGlyAlaHisIleAspCysGluValAlaProLeuProThrIleTyr 168  
 |||||  
 DB 552 GTCCAGAGGTGTTGGATCCACAAAGGCTCGAGAGTGGCCCATTCAGAGTGTCTCCA 611  
 QY 168 sArgGlnIysSerGluLeuSerAspGlyIleAlaMetLeuValAlaGlyAsnAspArg 188  
 |||||  
 DB 612 GGGACAAAGACTGACTGATTAATCTATCTCATCTGCTGCGGGGAAATGACCGCT 671  
 QY 188 IglAlaValIleThrGlnMetGlnGluValCysGlnThrIleGluAspAsnArgArg 208  
 |||||  
 DB 672 GCGACCATATCATCTAGTGGAGGATTCCTCCGTGAGTGCACCAAGAGAACTACAC 731  
 QY 208 gClnIysGlnLeuLeuAsnGln-ArgPheGlnThrIleCys-AlaValIleGluGluArg 227  
 |||||  
 DB 732 GGTAAAGGAAGAGCTAGCCACAGATTGACAGTTAGTATGCACTGATGATGAGAG 791  
 QY 228 LysGlyGluLeuLeuGlnAlaLeuAlaArgGlnGlnGlnIleGlnIleGlnIleGlnIle 247  
 |||||  
 DB 792 AAAAGTAGTGTGGTGCAGGATCACGCGAGGAGAGAGACAAAGCTTAAGTTCATCCGAG 851



QY 248 gylleuileargln 252  
 |||||||  
 Db 852 GGCCTCATCCAGCAG 866

RESULT 7  
 LOCUS BIT54358  
 DEFINITION 603026906F1 NIH\_MGC\_114 Homo sapiens cDNA clone IMAGE:5197560 5',  
 mRNA sequence.  
 ACCESSION BIT54358  
 VERSION BIT54358.1 GI:15745936  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
 REFERENCE 1 (bases 1 to 793)  
 AUTHORS NIH-MGC http://mhc.nci.nih.gov/  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cgapbs-remail.nih.gov  
 Tissue Procurement: Life Technologies, Inc.  
 cDNA Library Preparation: Life Technologies, Inc.  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov  
 Plate: LLM11494 row: k column: 01  
 High quality sequence stop: 793.

FEATURES  
 source  
 1..793  
 Location/Qualifiers  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5197560"  
 /clone\_1ib="NIH\_MGC\_114"  
 /lab\_host="DH10B"  
 /note="Organ: brain; Vector: pCMV-Sport6; Site:1: NotI;  
 Site-2: EcoRV (destroyed); RNA source anonymous pool of 6  
 male brains, age range 23-27 yo. library is oligo-dT  
 primed and directionally cloned (EcoRV site is destroyed  
 upon cloning). Average insert size 1.5 kb, insert size  
 range 1-3 kb. Library is normalized and enriched for  
 full-length clones and was constructed by C. Gruber  
 (Invtrogen). Research Genetics tracking code 019. Note:  
 this is a NIH-MGC Library."  
 BASE COUNT 201 a 208 c 228 g 156 t  
 ORIGIN

Alignment Scores:  
 Pred. No.: 5,14e-80 Length: 793  
 Score: 825.50 Matches: 153  
 Percent Similarity: 85.45% Conservative: 35  
 Best Local Similarity: 69.55% Mismatches: 28  
 Query Match: 43.17% Gaps: 2  
 Indels: 4  
 Gaps: 2

US-09-908-988b-2 (1-366) x BIT54358 (1-793)

QY 10 leuenglyasplahisasmetsaspasntleuglysglnleuilecysprollecys 29  
 |||  
 Db 140 CTGATTCAGGATGGGAAATCCCATGAGAACTTGGAGAGAGCTGATCTGCCCTATCTGC 199

QY 30 leuglumetpbeserlysprovalvalilleuaprocyglnhisasnlleucysarglys 49  
 |||||||  
 Db 200 CTGAGATGTTTACCAAGCCAGTGTCTGCGCGTCCAGCACACCTGTGCGGAG 259

QY 50 CysAlaAsnAspValPheGlnAlaSerAsnProLeuTrpGlnSerArgGlySerThrThr 69  
 |||||||  
 Db 260 TGTGCAATGACATCTTCCAGGCTGCAAAATCCCTACTGAGACAGCCGGGCGACGCTCAATG 319

QY 70 valserglyglyllyargpheaargcysprosercysargnhisgluvalvalleuasparg 89  
 |||||||  
 Db 320 TGCATGCTGGAGGCCGCTTTCGCGTCCGCCACCTGCCGACAGAGTGATCATGTGCTG 379

QY 90 hisglyvaltyrglyleuglnargasnlleuvalgluasnlleileaspilleylys 109  
 |||||||  
 Db 380 CACGAGTGTACGCCCTCAGAGAACTGCTGGTGAGAAATCATCATGATCTACAA 439

QY 110 glnglu---SerSerargproleuhsalalysalaglglnglnhisleuetycsglu 128  
 |||||||  
 Db 440 CAGAGTCTCCAGTCGCGCGCTG-----CAGAGGGCAGTACACCCCATGTGCAAGAG 493

QY 129 hisgluasprglyllyleasnleetyrcysleuetyrcysgluvalprothrcysserleu 148  
 |||||||  
 Db 494 CACGAAATGAGAAATCAATCAATCTACTCTCTCAGCTGTGAGAGTCCGCCACCTGCATG 553

QY 149 cyslysalpneghialahislyaspysgluvalalaproleuprothrlleylys 168  
 |||||||  
 Db 554 TCGAAGTGTGGGATTCACAGAGCCCTGCGAGTGGCCCTTTCAGAGTCTTCCAG 613

QY 169 arggllysserglyleuSeraspilyllealmetleuvalalagllyasnapargval 188  
 |||||||  
 Db 614 GGACAAAGACTGAACATGAATACGTATCTCATGCTGTGGCGGGAATGACCGTGG 673

QY 189 glalavallietnnglmetcglugluvalcysglnthrlleuglyasnapserarg-Arg 208  
 |||  
 Db 674 CAGACCATCATCTACCTGAGGATTCCTCGTCCAGTACCAAGAAACAGTACCCAG 733

QY 208 ggllysglnleuLeuasnnglnargphegluThrLeuCysAlaValleugluarg 227  
 |||  
 Db 734 GTACACGGAAGAGCTGAGCCAGAGTGTGACACGTTGATGCAATCCCGATGAGAGAG 791

RESULT 8  
 LOCUS BG761218  
 DEFINITION 602718564F1 NIH\_MGC\_49 Homo sapiens cDNA clone IMAGE:4858622 5',  
 mRNA sequence.  
 ACCESSION BG761218  
 VERSION BG761218.1 GI:14071871  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
 REFERENCE 1 (bases 1 to 1105)  
 AUTHORS NIH-MGC http://mhc.nci.nih.gov/  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cgapbs-remail.nih.gov  
 Tissue Procurement: ATCC/DCTD/DNP  
 cDNA Library Preparation: Ling Hong/Rubin Laboratory  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov  
 Plate: LLM1712 row: p column: 15  
 High quality sequence stop: 787.

FEATURES  
 source  
 1..1105  
 Location/Qualifiers  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:4858622"  
 /clone\_1ib="NIH\_MGC\_49"  
 /tissue\_type="melanotic melanoma, high MDR (cell line)"  
 /lab\_host="DH10B (phage-resistant)"  
 /note="Organ: skin; Vector: pOT87; Site:1: XhoI; Site:2:  
 EcoRI; cDNA made by oligo-dT priming. Directionally cloned  
 into EcoRI/XhoI sites using the following 5' adaptor:  
 GGCAAGAG(G). Size-selected >500bp for average insert size  
 1.8kb. Library constructed by Ling Hong in the laboratory  
 of Gerald M. Rubin (University of California, Berkeley)



/lab\_host="DH10B"  
 /note="Site\_1: SalI; Site\_2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN, Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5', GAGAGAGAGCGCGCCGACACCGCAGCTTTTCTTTTCTTTT 3'], cDNA was prepared by using triethanolamine-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. Second strand cDNA was prepared with the primer adapter of sequence [5', GAGAGAGAGATTCTGAGTATTAATTAATCCCGCCCCCCC 3']. cDNA was cleaved with BamHI and XhoI. Vector: a modified pBluescript KS(+) after bulk excision from Lambda PUC I."

BASE COUNT 183 a 178 c 179 g 153 t 1 others

## ALIGNMENT SCORES:

Pred. No.: 9.76e-79 Length: 694  
 Score: 813.00 Matches: 149  
 Percent Similarity: 89.84% Conservative: 19  
 Best Local Similarity: 79.68% Mismatches: 17  
 Query Match: 42.52% Indels: 2  
 DB: 10 Gaps: 2

US-09-908-988b-2 (1-366) x BB520266 (1-694)

QY 13 AspalahisAsnMetAspAsnLeuGluGlnLeuIleCysProIleCysLeuGluMet 32  
 DB 140 GAGAGACAGACACGATGAGTGGAAAGCAACGATCTGCCATCTGCTAGAGAT 199  
 QY 33 PheSerIysProValIleIleuProCysGlnHisAsnLeuCysArgIysCysAlaAsn 52  
 DB 200 TTGACGAGAGCCTGGTGCATCTCTCCCTGCCAGCAACCTGTGAGGAAATGTGCAGT 259  
 QY 53 AspyAlaheGlnAlaSerAsnProLeuTrpGlnSerArgIysSerThrValSerSer 72  
 DB 260 GACATCTTCAGGCGCTTAACCCGTAATACCAAGAGAGGACACACCGTGCATCA 319  
 QY 73 GtGtYtrpPheArgCysProSerCysArgHisGluValIleuAspArgHisGluVal 92  
 DB 320 GGGGGCGGCTTCCCTCTCCCTGCGACATAGAGTGTGTAGACAGATGGGCTC 379  
 QY 93 TyGtIleuGlnIleArgAsnLeuValGluAsnIleIleAspIleTyTrpGlnGlnSer 112  
 DB 380 TATGAGACTGCAGAGAGACGCTGCTGGAACATTAATGATATCTACAGCAGAAATCC 439  
 QY 113 SerArgProLeuHisAlaIleValGluGlnHisIleuMetCysGluGlnHisGluAspGlu 132  
 DB 440 ACCAGGCGCA--GAAAAAAATGAGCCAGCC--ATGTGTGAAGACATGAAAGAGAA 493  
 QY 133 LysIleAsnIleTyTrpCysLeuSerCysGluValProThrCysSerIleuCysValPhe 152  
 DB 494 CGCATCAACATCTATTCTCTGAAGTGTGAGAGTGCACCTGTCTCTGTGGAAGGTTT 553  
 QY 153 GtAlaHisIleLysAspCysGluValAlaProLeuProThrIleTyTrpArgIleLysSer 172  
 DB 554 GGGGCCCATTAAGAGTGCAGTGGGCTCCCTGATCATGTGTCCAGAGCACAAGTCA 613  
 QY 173 GluLeuSerAspGlyIleAlaMetLeuValAlaGlnAsnAspArgValGlnAlaValIle 192  
 DB 614 GACCTAGTGAAGTGTGTGTGTTGTGGAAGCAACGATAGATCCAGGAGGTGTATC 673  
 QY 193 ThrGlnMetGluGluValCys 199  
 DB 674 AGCCAGCTGAGAGACACCTGT 694

## RESULT 10

LOCUS BU033623 615 bp mRNA linear EST 07-DEC-2001  
 DEFINITION BU033623 NIBB Mochii normalized Xenopus neurula library Xenopus

ACCESSION laevis cDNA clone XL022F14 5', mRNA sequence.  
 VERSION BU033623  
 KEYWORDS GI:17414076  
 SOURCE EST.

## ORGANISM

African clawed frog.  
 Xenopus laevis

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## COMMENT

Expressed genes in X. laevis embryo  
 Unpublished (2001)  
 Contact: Tadasi Shin-i  
 Center For Genetic Resource Information  
 National Institute of Genetics  
 1111 Yata, Mishima, Shizuoka 411-8540, Japan  
 Tel: 81-559-81-6856  
 Fax: 81-559-81-6856  
 Email: tshini@genes.nig.ac.jp.  
 Location/Qualifiers

## FEATURES

## source

1..615  
 /organism="Xenopus laevis"  
 /db\_xref="taxon:8355"  
 /clone="XL022F14"  
 /clone\_lib="NIBB Mochii normalized Xenopus neurula library"  
 /tissue\_type="whole embryo"  
 /dev\_stage="stage 15"  
 /note="Vector: pBSR3; Site\_1: NotI; Site\_2: EcoRI; cDNAs were oligo-dT primed and directionally cloned. Staging according to Nieuwkoop and Faber. Library is subcloned and was constructed by N. Garrett and A.M. Zorn, (Wellcome/CRC Institute)."

## BASE COUNT

197 a 119 c 138 g 161 t

## ALIGNMENT SCORES:

Pred. No.: 7.39e-76 Length: 615  
 Score: 786.00 Matches: 145  
 Percent Similarity: 87.05% Conservative: 23  
 Best Local Similarity: 75.13% Mismatches: 23  
 Query Match: 41.11% Indels: 2  
 DB: 13 Gaps: 2

US-09-908-988b-2 (1-366) x BU033623 (1-615)

QY 1 MetAsnPheThrValGlyPheLysProLeuLeuGlyAspAlaHisAsnMetAspAsnLeu 20  
 DB 41 ATGAGCAGCTCATTTGCAAGTACAGTCAATTTCTCCAAAGAGAGCAAACTATGATTAATCTA 100  
 QY 21 GluLysGlnLeuIleCysProIleCysLeuGluMetPheSerIysProValIleIleu 40  
 DB 101 GAGAGCACTCATATATGTCATATGCTAGAAATGTTTCCAAAGCTGTGGCATATCTC 160  
 QY 41 ProCysGlnHisAsnLeuCysArgIysCysAlaAsnAspValPheGlnAlaSerAsnPro 60  
 DB 161 CCTGCGCACATTAATCTGTGTGAGAGTGTGCGAGATATATTCACAGCTTCCAAATACG 220  
 QY 61 LeuTrpGlnSerArgGlySerThrValSerSerGlyIleArgPheArgCysProSer 80  
 DB 221 TATTTGGCCAAACGAGAGAGAAACACAGTACGTCGGGTGAGAGTTTTCGCTCCATCG 280  
 QY 81 CysArgHisGluValIleuAspArgHisGluValTyTrpGlyLeuGlnIleArgAsnLeu 100  
 DB 281 TGTAGACATGAGTGTGTGTGACAGACATGCGGTATATGACTACAAAGAACTTACTG 340  
 QY 101 ValGluAsnIleIleAspIleTyTrpGlnGlnIleuSerSerArgProLeuHisAlaIle 120  
 DB 341 GTTGAATAATATTTGACCTTTATTAACGAAATCTACAAAGGCT--GAAAGGAAGACT 397  
 QY 121 GluGlnHisIleuMetCysGluGlnHisGluAspGlyIleAsnIleTyTrpCysLeuSer 140







University of Delaware

LOCUS  
DEFINITION

BASE COUNT	84 a	156 c	162 g	74 t	19 others
ORIGIN	from 5'-end"				
Alignment Scores:					
Pred. No.:	8,69e-67	Length:	495		
Score:	702.00	Matches:	130		
Percent Similarity:	87.82%	Conservative:	7		
Best Local Similarity:	83.33%	Mismatches:	18		
Query Match:	36,72%	Indels:	1		
DB:	13	Gaps:	0		
US-09-908-988B-2 (1-366) x BM488618 (1-495)					
QY	1	MetAsnPhetHrValGlyPheLysProLeuLeuGlyAspAlaHisAsnMetAsnLeu	20		
DB	29	ATGAACTTGGCGGTGGGGGCTGAAGCCGCTGTCTGGCGGAAGCGGAGCATGAGACCTTG	88		
QY	21	GIuLyGlnLeuIleCysProIleCysLeuGluMetPheSerLysProValValIleLeu	40		
DB	89	GAGAAACAGCTATCTGCCCCCATCTGCTGGAGATGTTCCACCAACCCGTCGTCTCTG	148		
QY	41	ProCysGlnHis-AsnLeuCysArgLysCysAlaAsnAspValPheGlnAlaSerAsnPr	60		
DB	149	CCCTCGACGACAAANNCTCTGCCGAAGCGCCCAACGACGCTTCCAGGCTCCCAACC	208		
QY	60	OleuTrpGlnSerArgLysSerThrThrValSerSerGluArgPheArgCysProSe	80		
DB	209	GCTGTGGCAGTGGCGGGGTTCGGCGCGGGGCCNMCCTGGGGGGCGNNNNNGTGGCCGTC	268		
QY	80	TCysArgHisGluValValLeuAspArgHisGluValTyrGluLeuGlnArgAsnLeu	100		
DB	269	GTGCCGCCACGAGGTGGTCTGTGACCGGCGACGGGGTGTACGGGGTGCACGGAACTGCT	328		
QY	100	ValAlaGlnAsnIleIleAspIleTyrLeuGlnIuSerSerArgProLeuHisAlaVal	120		
DB	329	GGTGGAGACATCATCGACATCTACAGAGGAAATCGCGCGCCCTGCACAGCCAAAGC	388		
QY	120	aGluGlnHisLeuMetCysGluGlnHisGluAspGluLysIleAsnIleTyrCysLeuSe	140		
DB	389	GGAGACGACGCTGATGTGCGAGGACACAGGNNACAGGCGCATCACTACTAGCTGCG	448		
QY	140	TCysGluValProThrCysSerLeuCysLysValPheGlnAlaHis	155		
DB	449	CTGGGNNNGCCCACTGCTCCCTTGCAGAGGTTCGGCGCTCAC	494		
RESULT 17					
LOCUS	CMS04GLR	1065 bp	DNA	linear	GSS 21-MAY-2000
DEFINITION	Tetradon nigroviridis genome survey sequence T7 end of clone				





## Alignment Scores:

Pred. NO.: 1.62e-65 Length: 1092  
 Score: 695.50 Matches: 160  
 Percent Similarity: 64.40% Conservative: 39  
 Best Local Similarity: 51.78% Mismatches: 90  
 Query Match: 36.38% Indels: 22  
 DB: 13 Gaps: 4

US-09-908-988b-2 (1-366) x B1114696 (1-1092)

```

QY 1 MetasnphethrValGlyPheIysProIeuLeuGlyAspAlaHisasnMetaspasnLeu 20
    |||...:|||||...:|||||...:|||||...:|||||...:|||||...:|||||...:|||||
Db 134 ATGAGCCGATCTCGAATTAACAATCTTTTCCAAAGACAGACAGACATGATGATTA 193
QY 21 GluIysGlnIleuIleCysProIleCysLeuGlnIleuMetPheSerIysProValIleu 40
    |||...:|||||...:|||||...:|||||...:|||||...:|||||...:|||||...:|||||
Db 194 GAGAGGACATCATCTGCTCCATCTGCTTAGAGATGTTCCAGAAACCTGTGTGATTC 253
QY 41 ProCysGlnHisasnLeuCysArgIysCysAlaAsnAspValPheGlnAlaSerAsnPro 60
    |||...:|||||...:|||||...:|||||...:|||||...:|||||...:|||||...:|||||
Db 254 CCTTGTAGCAACAACCTGTGTAGAAATGTGCGCATGATTTTCCAGCCCTTAACCG 313
QY 61 LeuTrpInserArgIysSerThrThrValSerSerGlyIysArgPheArgCysProSer 80
    |||...:|||||...:|||||...:|||||...:|||||...:|||||...:|||||...:|||||
Db 314 TATTTGCCACAAAGAGAGTACCAACCATGGCATCAGGGGCGCATTCGCTGCCATCC 373
QY 81 CysArgHisGlnValValIleuAspArgHisGlyValIysIleuGlnArgAsnLeu 100
    |||...:|||||...:|||||...:|||||...:|||||...:|||||...:|||||...:|||||
Db 374 TGTAGACATAGAGGTGTTGCGATAGACATGGGATATGTGACTTCAAGAGACCTGCTG 433
QY 101 ValGlu--AsnIleIleAspIleIleIleIleIleIleIleIleIleIleIleIleIle 120
    |||...:|||||...:|||||...:|||||...:|||||...:|||||...:|||||...:|||||
Db 434 GTGAGCAACTATCATTTGATGATTCACAGCGAGGTCCACAGCCCA-----GAAAGA 487
QY 120 IagIuGlnHisIleuMetCysGlnIleuHisGlnAspGlnIleuAsnIleIleIleIleIle 140
    |||...:|||||...:|||||...:|||||...:|||||...:|||||...:|||||...:|||||
Db 488 AATCCAGCAGCC-ATGTGGAGAGACATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 546
QY 140 erCysGlnVal--ProThrCysSerLeuCysIysValPheGlnAlaHisIysAspCysGln 159
    |||...:|||||...:|||||...:|||||...:|||||...:|||||...:|||||...:|||||
Db 547 ACTGCACAGTACCAACCTGCTCTGTGCAAGGTGA--GGTCACACAAAGACTGCCAG 605
QY 160 ValAla--ProIeuProThrIleIleIleIleIleIleIleIleIleIleIleIleIle 179
    |||...:|||||...:|||||...:|||||...:|||||...:|||||...:|||||...:|||||
Db 606 GGGCGTCACCATATCATCTGTACAGAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 665
QY 179 IametLeuValAlaGlyAsnAspArgValGlnAlaValIleThrGlnMetGlnIleVal 199
    |||...:|||||...:|||||...:|||||...:|||||...:|||||...:|||||...:|||||
Db 666 GCATACTCTGTGGGACAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 725
QY 199 YSGIlnThrIleGlnAspAsnSerArgArgGlnIleIleIleIleIleIleIleIleIle 219
    |||...:|||||...:|||||...:|||||...:|||||...:|||||...:|||||...:|||||
Db 726 GCAAACTATCGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 773
QY 219 hrLeuCysAlaValIleuGlnIleuArgIysGlyIleuLeuGlnAlaLeuAlaArgIle 239
    |||...:|||||...:|||||...:|||||...:|||||...:|||||...:|||||...:|||||
Db 774 CGGAGACAGCTGAGTACTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 833
QY 239 IlnGlnIleuIleuIleuArgValArgGlyLeuIleArgGlnIleIleIleIleIleIle 259
    |||...:|||||...:|||||...:|||||...:|||||...:|||||...:|||||...:|||||
Db 834 AGCCACACCCCAACCAACAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 884
QY 259 IysSerIleuIleuValGlnSerAlaIleGlnSerMetIleuIleuProGlnIleuIleu 279
    |||...:|||||...:|||||...:|||||...:|||||...:|||||...:|||||...:|||||
Db 885 AGGCAAGTCCGAGCTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 926
QY 279 yrLeuGlnAlaIleuIleuIleuIleuValGlnIleuIleuIleuIleuIleuIleu 299
    |||...:|||||...:|||||...:|||||...:|||||...:|||||...:|||||...:|||||
Db 927 ACAGAAACAGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 966
QY 299 IagIysArgProGlnPro 304
    |||...:|||||...:|||||...:|||||...:|||||...:|||||...:|||||...:|||||
  
```

```

Db 987 ACACCCACCCGAGACCA 1003
RESULT 19
B1114212
LOCUS
DEFINITION
B1114212
ACCESSION
B1114212
VERSION
B1114212.1 GI:14565113
KEYWORDS
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE
1 (bases 1 to 776)
NIH-MGC http://mgc.ncl.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Ling Hong/Rubin Laboratory
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM1840 row: P column: 18
High quality sequence stop: 650.
Location/Qualifiers
1..776
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5021825"
/clone_lib="NIH_MGC_17"
/tissue_type="rhodomyosarcoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: muscle; Vector: pOTB7; Site_1: EcoRI;
Site_2: XhoI; CDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(5). Size-selected >500bp
for average insert size 1.8kb. Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-CDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."
```

## BASE COUNT

226 a 210 c 190 g 150 t

## Alignment Scores:

Pred. NO.: 1.76e-65 Length: 776  
 Score: 693.00 Matches: 141  
 Percent Similarity: 77.57% Conservative: 25  
 Best Local Similarity: 65.89% Mismatches: 41  
 Query Match: 36.24% Indels: 7  
 DB: 13 Gaps: 1

US-09-908-988b-2 (1-366) x B1114212 (1-776)

```

QY 1 MetasnphethrValGlyPheIysProIeuLeuGlyAspAlaHisasnMetaspasnLeu 20
    |||...:|||||...:|||||...:|||||...:|||||...:|||||...:|||||...:|||||
Db 134 ATGAGCCGATCTCGAATTAACAATCTTTTCCAAAGAGACAGACATGATGATTA 193
QY 21 GluIysGlnIleuIleCysProIleCysLeuGlnIleuMetPheSerIysProValIleu 40
    |||...:|||||...:|||||...:|||||...:|||||...:|||||...:|||||...:|||||
Db 194 GAGAGGACATCATCTGCTCCATCTGCTTAGAGATGTTCCAGAAACCTGTGTGATTC 253
QY 41 ProCysGlnHisasnLeuCysArgIysCysAlaAsnAspValPheGlnAlaSerAsnPro 60
    |||...:|||||...:|||||...:|||||...:|||||...:|||||...:|||||...:|||||
Db 254 CCTTGTAGCAACAACCTGTGTAGAAATGTGCGCATGATTTTCCAGCCCTTAACCG 313
QY 61 LeuTrpInserArgIysSerThrThrValSerSerGlyIysArgPheArgCysProSer 80
    |||...:|||||...:|||||...:|||||...:|||||...:|||||...:|||||...:|||||
Db 314 TATTTGCCACAAAGAGAGTACCAACCATGGCATCAGGGGCGCATTCGCTGCCATCC 373
  
```

QY 81 CysargHisGluValValLeuAspArgHisGlyValTyrGlyLeuGlnArgAsnLeu 100  
 |||||  
 Db 374 TGTAGACATGAAGTGTCCGAGACATGGGATATGAGCTTACAGAGACCTGCTG 433  
 |||||  
 QY 101 ValGluAsnIleIleAspIleTyrGlyGlnGluSerSerArgProLeuHisAlaIysAla 120  
 |||||  
 Db 434 GTGGAATATATCATTTGACATCTACACAGACAGAGATCCACAGGCCA-----GAAAGAAA 487  
 |||||  
 QY 121 GluGlnHisLeuMetCysGluGlnHisGlu-AspGluIysIleAsnIleTyrCysLeu-S 140  
 |||||  
 Db 488 TCCGACACAGCCATGTGCGAGAACATGACAGAGACAGCCATCAACATCTACCTGCGAA 547  
 |||||  
 QY 140 eercyGluValProThrCysSerLeuGlyValPheGlyAla-HisLysAspCysGlu 159  
 |||||  
 Db 548 CCTGGAGATACCCACCATGCTCTGTGCGAGGTGTCAGAGTCCACCAAGAAGCTGCGAG 607  
 |||||  
 QY 160 ValAlaProLeuProThrIleTyrIlyArgGln-LysSerGluLeuSerAspGly-IleA 179  
 |||||  
 Db 608 GTGGCTCCCTCCTACATCATGTGACAGACAGAACAGTGTGACCTCGCGATGGCCATCG 667  
 |||||  
 QY 179 IamettLeuValAlaGlyAsnAspArgValGlnAlaValIleThrGlnMetGluValC 199  
 |||||  
 Db 668 CCACCCTCTGTGGGACAGACATCGCGTCGAGATGATCAGCAGCGTGGACACACAT 727  
 |||||  
 QY 199 ysgInThrIleGluAspAsnSerArgArgGln 209  
 |||||  
 Db 728 GGGACACATCGAGAAAGTGCAGAAAACAGA 759  
 |||||  
 RESULT 20 552 bp mRNA linear EST 11-DEC-2001  
 LOCUS BJ072360  
 DEFINITION BJ072360 NIBB Mochii normalized Xenopus tailbud library Xenopus  
 laevis cDNA clone X1098902 5', mRNA sequence.  
 ACCESSION BJ072360 GI:17502549  
 VERSION BJ072360.1  
 KEYWORDS EST.  
 SOURCE African clawed frog.  
 ORGANISM Xenopus laevis  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Amphibia; Batrachia; Anura; Mesobatrachia; Pipridae; Pipridae;  
 Xenopodinae; Xenopus.  
 1 (bases 1 to 552)  
 Kitayama, A., Terasaka, C., Mochii, M., Ueno, N., Shin-I, T. and Kohara  
 Y.  
 Expressed genes in X. laevis embryo  
 Unpublished (2001)  
 Contact: Tadasi Shin-I  
 Center For Genetic Resource Information  
 National Institute of Genetics  
 1111 Yata, Mishima, Shizuoka 411-8540, Japan  
 Tel: 81-559-81-6856  
 Fax: 81-559-81-6855  
 Email: tshin@genes.nig.ac.jp.  
 Location/Qualifiers  
 1..552  
 /organism="Xenopus laevis"  
 /db\_xref="taxon:8355"  
 /clone="X1098902"  
 /clone\_lib="NIBB Mochii normalized Xenopus tailbud  
 library"  
 /tissue\_type="whole embryo"  
 /dev\_stage="stage 25"  
 /note="Vector: pBSRN3; Site\_1: NotI; Site\_2: EcoRI; cDNAs  
 were oligo-dT primed and directionally cloned. Staging  
 according to Nieuwkoop and Faber. Library is subcloned  
 and was constructed by N. Garrett and A.M. Zorn,  
 (Wellcome/CRC Institute)."  
 BASE COUNT 174 a 109 c 123 g 145 t 1 others  
 ORIGIN  
 Alignment Scores: 4.e-65 Length: 552  
 Pred. No.:

Score: 687.00 Matches: 126  
 Percent Similarity: 86.14% Conservative: 17  
 Best Local Similarity: 75.90% Mismatches: 21  
 Query Match: 35.93% Indels: 2  
 DB: 13 Gaps: 2  
 US-09-908-988b-2 (1-366) x BJ072360 (1-552)  
 QY 1 MetaspRhetrValGlyPheIysProLeuLeuGlyAspAlaHisAsnMetaspAsnLeu 20  
 |||||  
 Db 61 ATGAGCACCCTATTTGACATGACATGCAATTCCTCCAAAGACAGCAACTTGTGATATCTA 120  
 |||||  
 QY 21 GluGlnGlnIleLeuGlyProIleCysLeuGlnMetPheSerLysProValIleLeu 40  
 |||||  
 Db 121 GAGAAACACATCATATGTGACATATGCGCTAGAAATGTTTCCAGGCTGTGGTCATATC 180  
 |||||  
 QY 41 ProCysGlnHisAsnLeuGlySerGlyCysAlaAsnAspValPheGlnAlaSerAsnPro 60  
 |||||  
 Db 181 CCTGCGCAGCATATATCTGTGTGGAAGTGTGCGAGATATATCCAGGCTTCCATATAC 240  
 |||||  
 QY 61 LeutpGlnSerArgGlySerThrValSerSerGlyIlyArgPheArgCysProser 80  
 |||||  
 Db 241 TATTTGCCAACACGAGAGAGAAACACTGTAGCATCGGGTGAGCTTTTGGCTCCATCG 300  
 |||||  
 QY 81 CysargHisGluValValLeuAspArgHisGlyValTyrGlyLeuGlnArgAsnLeu 100  
 |||||  
 Db 301 TGTAGACATGAAGTGTGTTGACAGACATGGCGGTATATGAGCTTACCAAGAACTTACTG 360  
 |||||  
 QY 101 ValGluAsnIleIleAspIleTyrGlyGlnGluSerSerArgProLeuHisAlaIysAla 120  
 |||||  
 Db 361 GTTGAATATATATGACCTTTATAACAGGAATCTACAGGCT--GAAAGGAAGACT 417  
 |||||  
 QY 121 GluGlnHisLeuMetCysGluGlnHisGluAspGluIysIleAsnIleTyrCysLeuSer 140  
 |||||  
 Db 418 GACCAACCA--ATGTGTATGATGACATGAGATGAGAAATATATTTATGTCTGAGAC 474  
 |||||  
 QY 141 CysGluValProThrCysSerLeuGlyValPheGlyAlaHisLysAspCysGluVal 160  
 |||||  
 Db 475 TGTGAAGTTCCTACCTGCTCCATGTGCAAAAGTGTGGAGCTCACAATAATGTGTGAGTT 534  
 |||||  
 QY 161 AlaProLeuProThrIle 166  
 |||||  
 Db 535 GCTCCTTTACACAAAGTT 552  
 |||||

Search completed: December 3, 2002, 14:46:29  
 Job time : 2074 secs

•  
•  
•  
•

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 3, 2002, 09:04:09 ; Search time 2053 Seconds  
(without alignments)  
11288.720 Million cell updates/sec

Title: US-09-908-988B-1  
Perfect score: 1431  
Sequence: 1 aagagatgtacagagatgt.....ataaagactcaagtcctcc 1431

Scoring table: IDENTITY-NUC  
Gapop 10.0, Gapect 1.0

Searched: 16154066 seqs, 8097743376 residues  
Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database:

TEST: 7  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estlin:\*  
4: em\_estmu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_hic:\*  
9: gb\_estcl:\*  
10: gb\_estl2:\*  
11: gb\_hic:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: gb\_gss:\*  
18: em\_gss\_hum:\*  
19: em\_gss\_hlv:\*  
20: em\_gss\_pln:\*  
21: em\_gss\_vrt:\*  
22: em\_gss\_fun:\*  
23: em\_gss\_mam:\*  
24: em\_gss\_mus:\*  
25: em\_gss\_other:\*  
26: em\_gss\_pro:\*  
27: em\_gss\_rod:\*

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	538.2	37.6	638	9	AA800245
C 2	508.2	35.5	608	12	BF284900
C 3	482.4	33.7	485	11	AK019655
C 4	443	31.0	555	12	BF855618
C 5	439.2	30.7	516	13	BM030114
C 6	420.8	29.4	461	9	AV006036

hits 1 & 6 are the  
and hits 2-5  
all the rest

C 7	416.2	29.1	512	12	BF855619
C 8	416.2	29.1	595	13	BF071661
C 9	413.6	28.9	432	10	BF732046
C 10	403.2	28.2	504	9	AT172700
C 11	361.2	25.2	392	9	AV006120
C 12	360	25.2	456	10	BE099923
C 13	350.2	24.5	912	13	BT758535
C 14	345.4	24.1	658	14	BM726338
C 15	342.6	23.9	793	13	BT754358
C 16	341.8	23.9	378	10	AW918753
C 17	341.4	23.9	440	10	BG375657
C 18	340.6	23.8	441	12	BG379943
C 19	333	23.3	868	12	BM486818
C 20	327.8	22.9	1105	12	BG761218
C 21	316	22.1	373	10	AW483766
C 22	312.8	21.9	694	10	BB520266
C 23	303.2	21.2	386	10	BE095943
C 24	289.2	20.2	316	10	BB140247
C 25	289	20.2	721	12	BG765856
C 26	284	19.8	790	12	BG674823
C 27	271.2	19.0	682	12	BF69996
C 28	260	18.2	613	13	BU033623
C 29	259.6	18.1	1161	12	BG163322
C 30	258.6	18.1	450	13	BM087922
C 31	258.4	18.1	983	12	BE205857
C 32	255.6	17.9	291	12	BF554230
C 33	249.4	17.4	1092	13	BT114696
C 34	246.8	17.2	475	9	AA428229
C 35	244.2	17.1	1081	12	BF309583
C 36	243.8	17.0	1065	17	CNS04GLR
C 37	240.2	16.8	420	9	AA447961
C 38	239.4	16.7	581	12	BF829678
C 39	239.2	16.7	329	9	AT763576
C 40	237.8	16.6	776	13	BT114212
C 41	237.6	16.6	672	10	BB662321
C 42	235.6	16.5	464	9	AT644642
C 43	234.6	16.4	573	9	AA434443
C 44	233.8	16.3	628	13	BU059953
C 45	233.8	16.3	628	13	BU059953

## ALIGNMENTS

RESULT 1  
AA800245/c 638 bp mRNA linear EST 30-APR-1998  
LOCUS EST189742 Normalized rat heart, Bento Soares Rattus sp. cDNA clone  
DEFINITION RHEM89 3' end, mRNA sequence.

ACCESSION  
AA800245  
VERSION  
AA800245.1 GI:2863200

KEYWORDS  
EST.

SOURCE  
Rattus sp.

ORGANISM  
Rattus sp.  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.

REFERENCE  
1 (bases 1 to 638)  
Lee, N.H., Glodok, A., Chandra, I., Mason, T.M., Quackenbush, J.,  
Kerlavage, A.R. and Adams, M.D.

AUTHORS  
Rat Genome Project: Generation of a Rat EST (RER) Catalog & Rat  
Gene Index

JOURNAL  
Unpublished (1998)

COMMENT  
The Institute for Genomic Research  
9712, Medical Center Drive, Rockville, MD 20850, USA

CONTACT  
Tel: (301)-838-3529  
Fax: (301)-838-0208  
Email: nhlee@tigr.org

FEATURES  
Seq primer: M13-21.  
Location/Qualifiers  
1..638  
/organism="Rattus sp."

```

/db_xref="ATCC (inhost):2007238"
/db_xref="taxon:10118"
/clone="RHFM89"
/clone_lib="Normalized rat heart, Bento Soares"
/Note="Organ: heart; Vector: pRT3Pac; Site_1: EcoRI;
Site_2: NotI"
BASE COUNT      105 a      221 c      165 g      147 t
ORIGIN
Query Match      37 6%; Score 538.2; DB 9; Length 638;
Best Local Similarity 91.7%; Pred. No. 2.4e-107;
Matches 595; Conservative 0; Mismatches 43; Indels 11; Gaps 2;

QY 783 GGAGAGGTTGTCACGACATTCAGAGCAACAGCCGACAGACCACTGTAAACCA 842
DB 638 GGAGAGGTTGTCACGACATTCAGAGCAACAGCCGACAGACCACTGTAAACCA 579
QY 843 GAGCTTCAGACACCTGTGCGCGTTTGGAGAGCGCAAGGGGCACTGTTCAAGCACT 902
DB 578 GAGCTTCAGACACCTGTGCGCGTTTGGAGAGCGCAAGGGGCACTGTTCAAGCGCT 519
QY 903 GCGCCGGAGAGAGAGAGAGTTGACAGCGCGTGGGGGCTTCATCCGACAGAGAGA 962
DB 518 GCGCTCGGTGCAAGAGAGAGAGAGTGTGAGCGCGTGTACCTCATCCGACAGTACGAGA 459
QY 963 CCACCTTGGAGGGCTTCCTCAAAAGCTGTGAGTCCGACATCCATCGAGAGAGCGCA 1022
DB 458 CCACCTTGGAGGGCTTCCTCAAAAGCTGTGAGTCCGACATCCATCGAGAGAGCGCA 399
QY 1023 GATGGCTCTACCTCCAGCAGAGCAAGAGAGTGTATCAACAAGTGGGGCAATGTGAA 1082
DB 398 GATGGCTCTACCTCCAGCAGAGCAAGAGAGTGTATCAACAAGTGGGGCAATGTGAA 339
QY 1083 GGTGAGACTGTGACAGAGCGCCGAGCCAGGCTATGAGAGCATGAGCAATCTGTGAG 1142
DB 338 GGTGAGACTGTGACAGAGCGCCGAGCCAGGCTATGAGAGCATGAGCAATCTGTGAT 279
QY 1143 CTTGAGACAGTGGCCCAATGTTTGGCAACATCGACTTCCAGCCGCGCGCTGGGA 1202
DB 278 CTTGAGACAGTGGCCCAATGTTTGGCAACATCGACTTCCAGCCGCGCGCTGGGA 219
QY 1203 TGAAGAGATGACACATGCTTGGATGGGAGAGGAGCAATGGCGGCTGGAGAGGA 1262
DB 218 TGAAGAGATGACACATGCTTGGATGGGAGAGGAGCAACATGGCGGCTGGAGAGGA 159
QY 1263 GCGGCTGACAGTGGCAGAGGCTCAGGCTGACAGTCCACTGTGATCCAGAGCGAC 1322
DB 158 GCGGCTGACAGTGGCAGAGGCTCAGGCTGACAGTCCACTGTGATCCAGAGCGAC 105
QY 1323 ACCGAGAGCGGAGCGCAAGGATGCTGAGATCTGCGAGAGAGCACCAGCCAGCAAGC 1382
DB 104 GCC-----CGGAGGCCAAGGAGATGCTGAGATCTGCGAGAGAGCACCAGCCAGCAATC 50
QY 1383 TCGGCTTCCGCGCCCGGAGAGTTCATTAAGAGACTCAAGTGTCC 1431
DB 49 TCGGCTTCCGCGCCCGGAGAGTTCATTAAGAGACTCAAGTGTCC 1

RESULT 2
BF284900/c      608 bp      mRNA      linear      EST 28-NOV-2000
LOCUS      BF284900
DEFINITION      EST149491 Rat Gene Index, normalized rat, Rattus norvegicus cdna
ACCESSION      BF284900
VERSION      BF284900.1 GI:11215970
KEYWORDS      EST.
SOURCE      Norway rat.
ORGANISM      Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE      1 (bases 1 to 608)
AUTHORS      Malek,R.L., Cho,J., Lee,Y., Karamycheva,S., Parvizli,B., Pertea,G.,

```

```

TITLE      Sultana,R., Tsai,J., White,J., Quackenbush,J. and Lee,N.H.
JOURNAL      Generation of ESTs from Normalized Rat Embryo, Bento Soares
COMMENT      Unpublished (2000)
              Other ESTs: EST350057
              Contact: Lee, NH
              The Institute for Genomic Research
              9712, Medical Center Drive, Rockville, MD 20850, USA
              Tel: (301)-838-3529
              Fax: (301)-838-0208
              Email: nhlee@tigr.org
              This clone is available through the ATCC, contact the ATCC
              tel#703-365-2700 for further information.
FEATURES
Source
1..608
/organism="Rattus norvegicus"
/db_xref="taxon:10116"
/clone="RG1EX57"
/clone_lib="Rat Gene Index, normalized rat, Rattus
norvegicus cdna"
/tissue_type="mixed tissue"
/lab_host="DH5-alpha"
/Note="Vector: pRT3Pac; Site_1: EcoRI; Site_2: NotI;
Combination of ROV, RBR, RRI, RLI, RLU, REM, RMU, RSP
RHE, RPC, RPN"
BASE COUNT      101 a      207 c      162 g      138 t
ORIGIN
Query Match      35.5%; Score 508.2; DB 12; Length 608;
Best Local Similarity 91.3%; Pred. No. 8.5e-101;
Matches 565; Conservative 0; Mismatches 43; Indels 11; Gaps 2;

QY 813 CAGCCGAGACAGAGACAGTGTAAACAGAGTTCAGAGACCTGTGCGGTTTGA 872
DB 608 CAGCCGAGACAGAGACAGTGTAAACAGAGTTCAGAGACCTGTGCGGTTTGA 549
QY 873 GGAGCGCAAGGGGCAATGCTTCAAGCAGTGGCCGGAGAGAGAGAGTGTACAGC 932
DB 548 GGAGCGCAAGGGGCAATGCTTCAAGCAGTGGCCGGAGAGAGAGTGTACAGC 489
QY 933 CTTGCGGGGCTTACCGCCAGTACGAGAGACCACTTGGAGGCTCTCAAGCTGTGGA 992
DB 488 CTTGCGGGGCTTACCGCCAGTACGAGAGACCACTTGGAGGCTCTCTAAGTGTGGA 429
QY 993 GTCCGCATCATGATGATGAGAGCGCGAGATGGCTCTACTCTCAGAGCAAGAGA 1052
DB 428 GTCCGCATCATGATGATGAGAGCGCGAGATGGCTCTACTCTCAGAGCAAGAGA 369
QY 1053 GCTGATCAACAAAGTGGGGCAATGTCGAAGTGTGAGAGTGGAGAGCGCCGAGCAGG 1112
DB 368 GCTGATCAACAAAGTGGGGCAATGTCGAAGTGTGAGAGTGGAGAGCGCCGAGCAGG 309
QY 1113 CTATGAGACATGAGACCAATTCCTGTGAGCTGTGAGACAGTGGCGCAATGTTGCAAC 1172
DB 308 CTATGAGACATGAGACCAATTCCTGTGAGCTGTGAGACAGTGGCGCAATGTTGCGGAC 249
QY 1173 CATGCACTTCCAGCCGCGCGCTGGGATGAAGAGATGAGACATGAGCTTGGATGG 1232
DB 248 CATGCACTTCCAGCCGCGCGCGCTGGGATGAAGAGATGAGACATGAGCTTGGACGG 189
QY 1233 GGAGAGGCGCAATGCGGCTGGAGAGAGCGGCTGAGAGTCCAGAAAGCTCAGGCT 1292
DB 188 GGAGAGGCGCAATGCGGCTGGAGAGAGCGGCTGAGAGTCCAGAAAGCTCAGAAAGTT-----T 135
QY 1293 GCACGTACCCAGTCTGATCCAGAGCGCACACCGAAGGGGAGGCCAAGGATGCTGAGG 1352
DB 134 GCACGTATCCGACTCTGATCCAGAGCGCACAGCC-----CGGAGGCCAAGGATGCTGAGG 80
QY 1353 ATCTGCGCAGAGACACCGCGCCAGCAAGCTGGCTTCCCGCGCGGAGAGGTTCTCA 1412
DB 79 ATCTGCGCAGAGACACCATGCGGTACCATCTCGGCTCCACCCCGAGGATGTTCTCA 20
QY 1413 TAAAGACTCAAGTGTCC 1431

```



Db 122 GTGCGGGGCAATGTCGAAGGTGAGAGCTGCGAGAGCGCCGAGCCAGCGCTATGAGAGCATG 181  
 QY 1126 GAGCAATTCCTGTGTGAGCGTGGAGCAGCGTGGCGGAATGTTGGAGACATCGACTTCCAG 1185  
 Db 182 GAGCAATTCCTGTGTGAGCGTGGAGCAGCGTGGCGGAATGTTGGAGACATCGACTTCCAG 241  
 QY 1186 CCGGCGCGCGCTGGGGATGAAGAGATGACGACATGCTTTGATGGAGGAGGCGCAT 1245  
 Db 242 CCGGCGCGCGCTGGGGATGAAGAGATGACGACATGCTTTGATGGAGGAGGCGCAT 301  
 QY 1246 GCGGGGCTGGAGAGAGCGGCTGGAGCGTCCAGAGAGCTCGACCTGACATCCCGAC 1305  
 Db 302 GCGGGGCTGGAGAGAGCGGCTGGAGCGTCCAGAGAGCTCGACCTGACATCCCGAC 361  
 QY 1306 TCTGATCCAGAGCGCAGACCCGAGCGGAGCCAGGATGCTGAGAGATGCTGGCGAGAGA 1365  
 Db 362 TCTGATCCAGAGCGCAGACCCGAGCGGAGCCAGGATGCTGAGAGATGCTGGCGAGAGA 421  
 QY 1366 CCACGCGCGCCAGCAAGCTTGGCTCCCGCCCGCGGAGAGTTCTCAATTAAGAGCTCAAG 1425  
 Db 422 CCACGCGCGCCAGCAAGCTTGGCTCCCGCCCGCGGAGAGTTCTCAATTAAGAGCTCAAG 481  
 QY 1426 TGTCT 1429  
 Db 482 TGCC 485

## RESULT 4

BF855618/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

COMMENT

FEATURES

SOURCE

555 bp mRNA linear EST 16-JAN-2001  
 RCO-FN0204-161100-031-d06 FN0204 Homo sapiens CDNA, mRNA sequence.  
 BF855618  
 BF855618  
 GI:12243362  
 EST.  
 human.  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 555)  
 Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,  
 Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,  
 Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bata,G.S., Simpson,D.H.,  
 Brunstein,A., deoliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare  
 M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and  
 Simpson,A.J.  
 Shotgun sequencing of the human transcriptome with ORF expressed  
 sequence tags  
 Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)  
 20202663  
 Contact: Simpson A.J.G.  
 Laboratory of Cancer Genetics  
 Ludwig Institute for Cancer Research  
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,  
 Brazil  
 Tel: +55-11-2704922  
 Fax: +55-11-2707001  
 Email: asimpson@ludwig.org.br  
 This sequence was derived from the PAPS/P-LICR Human Cancer Genome  
 Project. This entry can be seen in the following URL  
 (http://www.ludwig.org.br/scripts/gethtml2.pl?PL-RC0&t2-RC0-FN0204-  
 161100-031-d06&t3-2000-11-16&t4=1)  
 Seq primer: puc 18 forward  
 High quality sequence start: 51  
 High quality sequence stop: 120.  
 Location/Qualifiers  
 1..555  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone\_lib="FN0204"  
 /dev\_stage="Adult"  
 /note="Organ: prostate.normal; Vector: puc18; Site\_1: SmaI  
 ; Site\_2: SmaI; A mini-library was made by cloning  
 products derived from ORESTES PCR (U.S. Letters Patent

application No. 196,716 - Ludwig Institute for Cancer  
 Research) profiles into the pUC 18 vector. Reverse  
 transcription of tissue mRNA and cDNA amplification were  
 performed under low stringency conditions."

Query Match 31.0%, Score 443; DB 12; Length 555;  
 Best Local Similarity 91.3%; Pred. No. 1.5e-86;  
 Matches 470; Conservative 0; Mismatches 45; Indels 0; Gaps 0;

276 CCGCATGCTGGAGAGTGTCTCCAAAGCCGTGGTATCTTCCCTGCCAACAACACT 335  
 Db CCGCATGCTGGAGAGTGTCTCCAAAGCCGTGGTATCTTCCCTGCCAACAACACT 496  
 555 CCGCATGCTGGAGAGTGTCTCCAAAGCCGTGGTATCTTCCCTGCCAACAACACT 496  
 QY 336 GTGCGGCAAGTGTGTCCAAAGCGTCTCCAGGCGCTATCTTGTGGCATCCCGGG 395  
 Db 495 GTGCGGCAAGTGTGTCCAAAGCGTCTCCAGGCGCTATCTTGTGGCATCCCGGG 436  
 QY 396 CTCACAAAGTGTCTTCCAGAGAGAGTTCGATGCCATCTTGAAGCAGAGTTGT 455  
 Db 435 CTCACCAAGTGTCTTCCAGAGAGAGTTCGATGCCATCTTGAAGCAGAGTTGT 376  
 456 CTTGAGAGGAGTGTGTCTTCCAGGAGAGTTCGATGCCATCTTGAAGCAGAGTTGT 515  
 Db 375 CTTGAGAGGAGTGTGTCTTCCAGGAGAGTTCGATGCCATCTTGAAGCAGAGTTGT 316  
 QY 516 CATCTACAGAGAGTGTCTTCCAGGAGAGTTCGATGCCATCTTGAAGCAGAGTTGT 575  
 Db 315 CATCTACAGAGAGTGTCTTCCAGGAGAGTTCGATGCCATCTTGAAGCAGAGTTGT 256  
 QY 576 TGAGAGAGGAGAGAGAGAGATCAATCTACTGCTGAGCTGCGAGTGGCCACTGT 635  
 Db 255 CGAGAGAGGAGAGAGAGAGATCAATCTACTGCTGAGCTGCGAGTGGCCACTGT 196  
 636 CTCTCTCTGCAAGTGTGTGCGCGCCCAAGAGAGTGTGAGTGGCCCTGCGCCACCAT 695  
 Db 195 CTCTCTCTGCAAGTGTGTGCGCGCCCAAGAGAGTGTGAGTGGCCCTGCGCCACCAT 136  
 QY 696 TTCAACAAAGGAGAGAGAGTGTGAGTGGAGTGGATGGATGGTGTGGGGGCAATGA 755  
 Db 135 TTCAACAAAGGAGAGAGAGTGTGAGTGGAGTGGATGGATGGTGTGGGGGCAATGA 76  
 QY 756 CCGTGTGAGAGAGAGTGTGAGTGGAGTGGATGGATGGTGTGGGGGCAATGA 790  
 Db 75 CCGTGTGAGAGAGAGTGTGAGTGGAGTGGATGGATGGTGTGGGGGCAATGA 41

RESULT 5

BM030114

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

COMMENT

516 bp mRNA linear EST 05-NOV-2001  
 BM030114  
 488634 MARC 280V Bos taurus CDNA 5', mRNA sequence.  
 BM030114  
 GI:16743684  
 EST.  
 cow.  
 Bos taurus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 Bovidae; Bovinae; Bos.  
 1 (bases 1 to 516)  
 Smith,T.P.L., Grosse,W.M., Freking,B.A., Roberts,A.J., Stone,R.T.,  
 Casas,E., Wray,J.E., White,J., Cho,J., Fahrnkung,S.C., Bennett  
 G.L., Heaton,M.P., Laegreid,W.W., Rohrer,G.A., Chitko-McKown,C.G.,  
 Pettea,G., Holt,I., Karamycheva,S., Liang,F., Quackenbush,J. and  
 Keele,J.W.  
 Sequence evaluation of four pooled-tissue normalized bovine cDNA  
 libraries and construction of a gene index for cattle  
 Genome Res. 11 (4), 626-630 (2001)  
 21180013  
 Contact: Smith TPL  
 USDA, ARS, US Meat Animal Research Center  
 PO Box 166, Clay Center, NE 68933-0166, USA





Db 239 GGATGACGACATGGCTTGTGATGGGAGGAGGGGCAATGCGGGGCTGGAGAGAGAGCGCT 298  
 Oy 1269 GGAGGTGCGAGAGGCTCAGGCTGCTGACTGACCCGACTGTGATCCAGAGGCAACCCGA 1328  
 Db 299 GACGTGCGAGAGGCTCAGGCTGCTGACTGACCCGACTGTGATCCAGAGGCAACCCGA 358  
 Oy 1329 AGCGGAGCGAAGGATGCTGAGATCTGCGAGAGACCAAGCGGCGACCAAGCTGGCT 1388  
 Db 359 AGCGGAGCGAAGGATGCTGAGATCTGCGAGAGACCAAGCGGCGACCAAGCTGGCT 418  
 Oy 1389 TCCGCGCGCGGAGAGGTTCTCAATTAAGGACTCAAGTGTCC 1431  
 Db 419 TCCGCGCGCGGAGAGGTTCTCAATTAAGGACTCAAGTGTCC 461

RESULT 7  
 LOCUS BF855619 512 bp mRNA linear EST 16-JAN-2001  
 DEFINITION RC0-FN0204-161100-031-e06 FN0204 Homo sapiens cDNA, mRNA sequence.  
 ACCESSION BF855619  
 VERSION BF855619.1 GI:12243363  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens

REFERENCE  
 AUTHORS Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 512)

DIAS NETO, E., GARCIA CORREA, R., VEJJOVSKI-ALMEIDA, S., BRIONES, M. R., NEAL, M. A., DA SILVA, W. Jr., ZAGO, M. A., BORDIN, S., COSTA, F. F., GOLDMAN, G. H., CARVALHO, A. F., MATSUKUMA, A., BALA, G. S., SIMPSON, D. H., BRUNSTEIN, A., DEOLIVEIRA, P. S., BUCHER, P., JONGENEEL, C. V., O'HARE, M. J., SOARES, F., BRENTANI, R. R., REIS, L. F., DE SOUZA, S. J., and SIMPSON, A. J.

TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags  
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)  
 MEDLINE 20202663  
 COMMENT Contact: Simpson A.J.G.  
 Laboratory of Cancer Genetics  
 Ludwig Institute for Cancer Research  
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil  
 Tel: +55-11-2704922  
 Fax: +55-11-2707001  
 Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL  
 (http://www.ludwig.org.br/scripts/gethtml2.pl?l=RC0&t=2-RC0-FN0204-161100-031-e06&t=3-2000-11-16&t=1)  
 Seg primer: puc 18 forward  
 High quality sequence start: 42  
 High quality sequence stop: 512.

FEATURES  
 source Location/Qualifiers  
 1..512

/organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone\_lib="FN0204"  
 /dev\_stage="Adult"  
 /note="Organ: prostate\_normal; Vector: puc18; Site\_1: SmaI; Site\_2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent Application No. 196/716 - Ludwig Institute for Cancer Research) profiles into the puc18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

BASE COUNT 104 a 138 c 157 g 113 t

ORIGIN

Query Match 29.1%; Score 416.2; DB 12; Length 512;  
 Best Local Similarity 91.0%; Pred. No. 1e-80;  
 Matches 465; Conservative 0; Mismatches 43; Indels 3; Gaps 2;

Oy 276 CCCCATCTGCTGAGATGTTCTCCAAAGCCGTGGATCTTGCCCTGCCAACACACCT 335  
 Db 512 CCCCATCTGCTGAGATGTTCTCCAAAGCCGTGGATCTTGCCCTGCCAACACACCT 453  
 Oy 336 GTGCCGCAAGTGTGCCAACGACGCTTCCAGGCTCTTAATCTGTGTGGCAATGCCGGG 395  
 Db 452 GTGCCGCAAGTGTGCCAACGACGCTTCCAGGCTCTTAATCTGTGTGGCAATGCCGGG 393  
 Oy 396 CTCCACAAAGGTGCTCTCAGAGAGAGGTTTCGATGCCCATCTGTAGGACAGAGTGT 455  
 Db 392 CTCCACAAAGGTGCTCTCAGAGAGAGGTTTCGATGCCCATCTGTAGGACAGAGTGT 333  
 Oy 456 CCTGAGACAGCATGCTGTCTATGCGCTGACGCGAAGCTGTGATGAGAACATATTGA 515  
 Db 332 CTTGAGACAGCATGCTGTCTATGCGCTGACGCGAAGCTGTGATGAGAACATATTGA 273  
 Oy 516 CATCTACAGAGAGGATGCTCTCAGAGAGAGGTTTCGATGCCCATCTGTAGGACAGAGTGT 575  
 Db 272 CATCTACAGAGAGGATGCTCTCAGAGAGAGGTTTCGATGCCCATCTGTAGGACAGAGTGT 213  
 Oy 576 TGAGAGACAGAGAGGATGATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 635  
 Db 212 CGAGAGACATGAGAGAGAGATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 153  
 Oy 636 CTCTCTCTGAGAGTGTTCGCGCCGACAGAGCTGTGAGGTGCGCCCTGTGCCACCAT 695  
 Db 152 CTCTCTCTGAGAGTGTTCGCGCCGACAGAGCTGTGAGGTGCGCCCTGTGCCACCAT 93  
 Oy 696 TTACAACAG--CCAGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 753  
 Db 92 TTACTTAACGCGCCCGAAGATGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 33  
 Oy 754 GACCGGTGTGAGGCA-GTATACACCCAGATG 783  
 Db 32 GACCGGTGTGAGGCA-GTATACACCCAGATG 2

RESULT 8  
 LOCUS BJ071661 595 bp mRNA linear EST 11-DEC-2001  
 DEFINITION BJ071661 NIBB Mochii normalized Xenopus tailbud library Xenopus laevis cDNA clone X1095D10 5', mRNA sequence.  
 ACCESSION BJ071661  
 VERSION BJ071661.1 GI:17501850  
 KEYWORDS EST.  
 SOURCE African clawed frog.  
 ORGANISM Xenopus laevis

REFERENCE  
 AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae; Xenopodinae; Xenopus. 1 (bases 1 to 595)

TITLE  
 JOURNAL  
 COMMENT Kitayama, A., Teresaka, C., Mochii, M., Ueno, N., Shin-I, T. and Kohara Y.  
 Expressed genes in X. laevis embryo  
 Unpublished (2001)  
 Contact: Tadasu Shin-I  
 Center For Genetic Resource Information  
 National Institute of Genetics  
 111 Yata, Mishima, Shizuoka 411-8540, Japan  
 Tel: 81-559-81-6856  
 Fax: 81-559-81-6855  
 Email: tshin@genes.nig.ac.jp.

FEATURES  
 source Location/Qualifiers  
 1..595

/organism="Xenopus laevis"  
 /db\_xref="taxon:8355"  
 /clone\_lib="X1095D10"  
 /clone\_id="NIBB Mochii normalized Xenopus tailbud library"  
 /tissue\_type="whole embryo"  
 /dev\_stage="stage 25"  
 /note="Vector: pBSRN3; Site\_1: NotI; Site\_2: EcoRI; CDNAS were oligo-dT primed and directionally cloned. Staging



Db 61 TCACAGAGTCGGGGCAATGTCGAAGTGGAGCTGGCAGAGCCCGGACCAAGCTATG 120  
 QY 1118 AGAGCATGAGCAATTTCTGTGAGCGTGGAGCAGCTGGCCGAATGTTGCGAACCATCG 1177  
 Db 121 AGACCATGAGCAATTTCTGTGAGCGTGGAGCAGCTGGCCGAATGTTGCGAACCATCG 180  
 QY 1178 ACTTCACGCGGCGCGCTGGATGAGAGAGATGACGATGCTTGGATGGGAGG 1237  
 Db 181 ACTTCACGCGGCGCGCTGGATGAGAGAGATGACGATGCTTGGATGGGAGG 240  
 QY 1238 AGGCAATGCGGGCTGGAGAGAGCGCTGACGTGCGAGAGGCTCAGGCTGACT 1297  
 Db 241 AGGCAATGCGGGCTGGAGAGAGCGCTGACGTGCGAGAGGCTCAGGCTGACT 300  
 QY 1298 GACCGACTGATGATCGACAGCGCACACCCGGAAGCGGAGCCAGAGGATGCTGAGATCTG 1357  
 Db 301 GACCGACTGATGATCGACAGCGCACACCCGGAAGCGGAGCCAGAGGATGCTGAGATCTG 360  
 QY 1358 CGCAGAGACACCGCGCCACCAAGCTCGGCTCCCGCCCGGGAAGTTCTCAATAAAG 1417  
 Db 361 CGCAGAGACACCGCGCCACCAAGCTCGGCTCCCGCCCGGGAAGTTCTCAATAAAG 420  
 QY 1418 GACTCAAGTGC 1429  
 Db 421 GACTCAAGTGC 432

RESULT 10  
 AT12700/c 504 bp mRNA linear EST 08-JUN-1999  
 LOCUS AT12700  
 DEFINITION UI-R-AFL1-aa-e-02-0-UI.s1 UI-R-AFL1 Rattus norvegicus CDNA clone  
 VERSION AT12700  
 KEYWORDS AT12700.1 GI:5016500  
 SOURCE Norway rat.  
 ORGANISM Rattus norvegicus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
 Rattus.  
 1 (bases 1 to 504)  
 Bonaudo, M.F., Lennon, G. and Soares, M.B.  
 Normalization and subtraction: two approaches to facilitate gene  
 discovery  
 Genome Res. 6 (9), 791-806 (1996)  
 97044477  
 CONTACT: Soares, MB  
 PROGRAM: Rat Gene Discovery and Mapping  
 UNIVERSITY: University of Iowa  
 451 Eckstein Medical Research Building Iowa City, IA 52242, USA  
 Tel: 319 335 8250  
 Fax: 319 335 9565  
 Email: msoares@blue.weeg.uiowa.edu  
 The sequence contained an oligo-dT track that was present in the  
 oligonucleotide that was used to prime the synthesis of first  
 strand cDNA and therefore this may represent a Donatide poly A  
 tail. The sequence tag present in the cDNA between the NotI site  
 and the oligo-dT track served to verify it as a clone from the  
 normalized AV canal at 15 dpc library CDNA library Preparation:  
 M.B. Soares Lab Clone distribution: clones will be available  
 through Research Genetics (www.resgen.com)  
 Seq primer: M13 Forward  
 POLY-A: yes.

# FEATURES

1. 504  
 Location/Qualifiers  
 /organism="Rattus norvegicus"  
 /strain="Sprague-Dawley"  
 /db\_xref="taxon:10116"  
 /clone="UI-R-AFL1-aa-e-02-0-UI"  
 /clone\_11b="UI-R-AFL1"  
 /dev\_stage="adult"  
 /lab\_host="DH10B (Life Technologies)"  
 /note="Vector: pT7T3D-Pac (Pharmacia) with a modified

polylinker. Site 1: Not I; Site 2: Eco RI; The UI-R-AFL1  
 library is a normalized library constructed from 15 dpc  
 rat atrioventricular (AV) canal. The tag is a string of 5  
 nucleotides present between the Not I site and the  
 oligo-dT track. The library was constructed as described  
 by Bonaudo, Lennon and Soares, Genome Research 6: 791-806  
 , 1996. Tissue provided by Jim Lin, Department of Biology,  
 University of Iowa.  
 TAG-UI-R-AFL1  
 TAG-TISSUE=AV canal at 15 dpc  
 TAG-SDO-GAAGC

Query Match 28.2%; Score 403.2; DB: 9; Length 504;  
 Best Local Similarity 91.2%; Pred. No. 7,1e-76;  
 Matches 454; Conservative 0; Mismatches 33; Indels 11; Gaps 2;

QY 934 GTGCGGGGCTCATCCGCCAGTACGAGACCACTTGGAGGCTCTCAAGAGTGGAG 993  
 Db 504 GTGCGTACCTCATCCGCCAGTACGAGACCACTTGGAGGCTCTCAAGAGTGGAG 445  
 QY 994 TCCGCCATCCAGTATGAGGAGCCGAGATGGCTCTCTACCTCCAGCAGCAAGAG 1053  
 Db 444 TCCGCCATCCAGTATGAGGAGCCGAGATGGCTCTCTACCTCCAGCAGCAAGAG 385  
 QY 1054 CTGATCAACAAGTTCGGGGCAATGTCGAAGTGGAGCTGGCAGAGACCGCGAGCAGC 1113  
 Db 384 CTGATCAACAAGTTCGGGGCAATGTCGAAGTGGAGCTGGCAGAGACCGCGAGCAGC 325  
 QY 1114 TATGAGAGCATGAGACATTTCTGTGAGCTGGAGCAGCAGTGCCGAATGTCGAAC 1173  
 Db 324 TATGAGAGCATGAGACATTTCTGTGAGCTGGAGCAGCAGTGCCGAATGTCGAAC 265  
 QY 1174 ATGCACTTCAGCGCGGCGCGCTGGGATGAAGAGATGACGACATGCTTGGATGG 1233  
 Db 264 ATGCACTTCAGCGCGGCGCGCTGGGATGAAGAGATGACGAGTGTGACTTGGACGG 205  
 QY 1234 GAGGAGGCAATCGGGGCTGAGAGAGAGCGGCTGGAGCTGCCAGAGCTCAGGCTG 1293  
 Db 204 GAGGAGGCAATCGGGGCTGAGAGAGAGCGGCTGGAGCTGCCAGAGCTCAGGCTG 151  
 QY 1294 CACTGACCGGACTGTATCCAGAGCGCACGCCGGAAGCGGAGGAGATGCTGAGGA 1353  
 Db 150 CACTGATCCGACTGTATCCAGAGCGCACGCCGGAAGCGGAGGATGCTGAGGA 96  
 QY 1354 TCTGCGCAGAGACACCGCGCCACCAAGCTGGCTTCCCGCCCGGGAAGGTTCTCAAT 1413  
 Db 95 TCTGCGCAGAGACACCGCGCTTACCATTCGCGCTCCACCCAGGAGATGTTCTCAAT 36  
 QY 1414 AAAGGACTCAAGTGTCC 1431  
 Db 35 AAAGGACTCAAGTGTCC 18

RESULT 11  
 AV006120 392 bp mRNA linear EST 25-AUG-1999  
 LOCUS AV006120  
 DEFINITION AV006120 Mus musculus C57BL/6J heart Mus musculus CDNA clone  
 VERSION AV006120  
 KEYWORDS AV006120.1 GI:4782970  
 SOURCE EST.  
 ORGANISM house mouse.  
 Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE  
 AUTHORS  
 1 (bases 1 to 392)  
 Akahira, S., Akiyama, J., Fukuda, S., Fukunishi, Y., Funayama, T., Hara  
 , A., Hayatsu, N., Horii, F., Ishikawa, T., Itoh, M., Iwama, M., Kawai, J.,  
 Kikuchi, N., Kojima, Y., Matsuyama, T., Nitsuma, H., Oda, H., Owa, C.,  
 Sato, K., Shibata, Y., Shigemoto, Y., Shiraki, T., Sogabe, Y., Sugahara



QY	1100	GGCCGAGCCAGGCATATGAGCATGGACCAATTCCTCTGAGCCGTGAGACACAGCTGGCCG	1159
Db	336	GGCCGAGCCAGGCATATGAGCATGGACCAATTCCTCTGTGATCGGGAGCACAGTGGCCG	277
QY	1160	AAATGTTCGGAACATCGACTTCCAGCCGGGCGCCGTGGGGATGAAGAGATGACGACA	1219
Db	276	AAATGTTCGGGACCATCGACTTCCAGCCGGGCGCCCTCTGGGATGAAGAGATGACGAGG	217
QY	1220	TGGCTTTGGATGGGAGGAGGAGGCATGCGGGGCTGAGAGGAGGAGGCGCTGGACGTGGCAG	1279
Db	216	TGACTTTGGAGCGGGAGAGGGCAACAGGGGCTTGAGAGGAGGAGCGGCTTGGACGGGCCAG	157
QY	1280	AAGGCTCAGGCGCTGCACATGACCCGACTCTGATTCACAGAGCGCACACCCGAGCGGAGCCA	1339
Db	156	AAGGTT-----TGACTGTATCGACTCTGATTCACAGAGCGCACGCCG----CGGAGGCCA	108
QY	1340	AGGATGCTGAGAGATCTGCGCAGAGACCAACCGCGCCACCAAGCTCAGGGTTCCCGGCCCGG	1399
Db	107	AGGGATGCTGAGAGATCTGCGCAGAGACCAACCGCGCTACCCATCTCGGCTCCACCCCGAG	48
QY	1400	GGAAGTCTTCATTAAGAGACTCAAGTGTTC	1429
Db	47	GGATGGTCTTCATTAAGAACTCATGTGTTC	18

[illegible]

REFERENCE AUTHORS TITLE JOURNAL COMMENT
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo. 1 (bases 1 to 912) NIH-MGC <a href="http://mgc.nci.nih.gov/">http://mgc.nci.nih.gov/</a> National Institutes of Health, Mammalian Gene Collection (MGC) unpublished (1999) Contact: Robert Strausberg Ph.D.

tissue procurement: Life Technologies, Inc.  
 cDNA library Preparation: Life Technologies, Inc.  
 CDNA library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MCC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
 Plate: LHAM1483 row: j column: 15  
 High quality sequence start: 25  
 High quality sequence stop: 890.  
 Location/Qualifiers  
     1..912  
 source

FEATURES	SOURCE
Location/Qualifiers	
1..912	
/organism="Homo sapiens"	
/db_xref="taxon.9606"	
/clone="IMAGE:5193326"	
/clone_id="NIH_MGC_114"	
/lab_host="DH10B"	
/note="Organ: brain; Vector: PCMV-SPORT6; Site.1: NotI; Site.2: EcoRV (destroyed); RNA source anonymous pool of 6 male brains, age range 23-27 yo. Library is oligo-dt primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.5 kb, insert size range 1-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (University of California, San Diego). Research (this is a NIH_MGC library." (Invitrogen)	
236 a	232 c 266 g 178 t

Query Match	24.5%;	Score 350.2;	DB 13;	Length 912;
Best Local Similarity	70.1%;	Pred. No. 3.6e-66;		
Matches 515; Conservative	0;	Mismatches 213;	Indels 7;	Gaps 3

OY	226	CTGCTTGGGGGATGGCGCAACAACATGGGCAACTTTGGAGAACACAGCTCAATTGGCCCATCTGC	285
Db	167	CTGATCCAGGATGGGATGGAATCCCATGGGAACTTGGAGAACACAGCTGATCTGCGCTTATCTGC	226
OY	286	CTGGAGATGTTTCCAAAGCCCGTGGGATCTTGACCTTGCCCTGGCAACACAACTGTGGCCGACG	345
Db	227	CTGGAGATGTTTACCAAGCCAGTGGATCTTGCCCTGGCAGCACAACTGTGGCCGAGG	286
OY	346	TCTGCCACGACAGTCTTCCAGGCGTCTAATCTCTGTGGCAATCCCGGGGCTCCACACAG	405
Db	287	TCTGCCAAAGACATCTTCCAGGCTGTGAAATCCCTACTGGACACAGCGGGGAGCTCAGTG	346
OY	406	GTGTCTTCGAGGAGAGTTCGATGGCCCATTTGTGAGGACGAGAGTTTCTCTGGACAG	465
Db	347	TCCATGTCTGGAGGCGCGTTTCCGCTGCCCCACCTGCGCCACGAGGATGATCGATCT	406
OY	466	CATGCTGTCTATGGCCTGCGAGGCAACCTGCTAGTGGAGAACATCATTTACATCTACAG	525
Db	407	CACGAGTGTACGGCCTGCGAGGAACTGCTGGTGGAGAACATCATGCACTTACAAA	466
OY	526	CAGGAGTCTCTCCCGGCCACTGCAAGCCCAAAGCTGAACAGCACCTCATGTGTGAGGAGAC	585
Db	467	CAGGAGTCTCCAGTGGCGCCGCTGCAGAAAG--GGCAGTCAACCCCATGTGCAAGGAGAC	523
OY	586	GAGGACGAGAGATCAACATCTACTGCGCTGAGCTGCGAGTGGGAGGCCCCACCTGCTCTCTGC	645
Db	524	GAAGTGTAGAAATTAACATCTACTGCTTACAGTGTGAAGTGGCCCACTGCTCCATGTGC	583
OY	646	AAGTTTTTCGCGCCCCACAGAGACTGTGAGTGGCCCTCTGCCCACTTTTACAAAGC	705
Db	584	AAGGTGTTGGGATCCACAAAGCCTGGCAGGAGGCCCATTTGCAAGTGTCTCCAGGA	643
OY	706	CAGAAAGTGTAGCTGAGCGATGGCATGCCGATGCTGTGTGGCGGGCAATGACCGTGTGCAG	765
Db	644	CAAAAGACTGGAACGTAAATACGTGATCTCCATGCTGTGGGGGGAATGACCGTGTGCAG	703
OY	766	GCAGATGACCCCGAGATGAGGAGAGGTGGCCACAGCACTTGGAGA--CAACACCCGAGACA	824
Db	704	ACCATCATCACTCAAGCTGAGAGATTTCCCGTGAAGTGAACCAAGGAAACAATCTCACAGT	763
OY	825	GAAGCAACTGTTAAACAGAGAGTTTCAGACCCCTGTGCGCGGTTTTTGGAGAGCCCAAGG	884
Db	764	AAAGGAGAGAGCTGAGCCGAGAGTTTGAACACTTGTATCCATCTCTGGATGAGAGAAAG	823
OY	885	CGAAGTGGTTAAGCACTGGCCCCGGGAGCAGAGAGAGAAGTTTGCAGCGCGTGGCGGCTT	944
Db	824	TGAATTTGCTGCAAGGATCAACGACAGGAGCGGAGAAAAAGTTAGTTCT--ATCAGAGCCT	880
OY	945	CATCCGCCAGTACGG	959
Db	881	CATCCAGCAGTACAG	895

RESULT 14	
BM726338	
LOCUS	BM726338 658 bp mRNA linear EST 01-MAR-2002
DEFINITION	UI-E-EJ0-alh-n-03-0-UI.r1 UI-E-EJ0 Homo sapiens cDNA clone
ACCESSION	UI-E-EJ0-alh-n-03-0-UI 5', mRNA sequence.
VERSION	BM726338
KEYWORDS	BM726338.1 GI:19047671 EST.
SOURCE	human.
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo. 1 (bases 1 to 658)
AUTHORS	Bonaldo,M.F., Lennon,G. and Soares,M.B.
TITLE	Normalisation and subtraction: two approaches to facilitate gene discovery



JOURNAL  
MEDLINE  
COMMENT

Genome Res. 6 (3), 791-806 (1996)  
9704477  
Contact: Soares, MB  
Program for Rat Gene Discovery and Mapping  
University of Iowa  
451 Eckstein Medical Research Building Iowa City, IA 52242, USA  
Tel: 319 335 8250  
Fax: 319 335 9565  
Email: msoares@blue.weeg.uiowa.edu  
Tissue Procurement: Dr. Gregg Hageman  
cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa  
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa  
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
Clone Distribution: Researchers may obtain clones from Research  
Genetics (www.resgen.com).  
Seq primer: M13 Reverse.

# FEATURES

source

1. .658  
Location/Qualifiers

/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="UI-E-EJ0-ah-n-03-0-UI"  
/clone\_lib="UI-E-EJ0"  
/tissue\_type="fetal eyes, lens, eye anterior segment,  
optic nerve, retina, Retina Foveal and Macular, RPE and  
choroid"  
/dev\_stage="fetal and adult"  
/lab\_host="DH10B (Life Technologies) (T1 phage resistant)"  
/note="Organ: eye; Vector: pT73-Pac (Pharmacia) with a  
modified polylinker; Site\_1: EcoR I; Site\_2: Not I;  
UI-E-EJ0 is a subtracted cDNA library constructed  
according to Bonaldo, Lennon and Soares, Genome Research,  
6:791-806, 1996. First strand cDNA synthesis was primed  
with an oligo-dT primer containing a Not I site. Double  
stranded cDNA was ligated to an EcoR I adaptor, digested  
with Not I, and cloned directionally into pT73-Pac  
vector. The oligonucleotide used to prime the synthesis of  
first-strand cDNA contains a library tag sequence that is  
located between the Not I site and the (dT)18 tail. The  
sequence tags for this library are: fetal eyes, AGAATCAAGA  
; lens, CGATTAGCGA; eye anterior segment, AATGCCGAT;  
optic nerve, CCATTAGTGT; retina, CCGCG; Retina Foveal and  
Macular, GTCC; RPE and Choroid, ACCCTA. This library was  
created for the program, Gene Discovery in the Visual  
System, supported by National Eye Institute (NEI)."

BASE COUNT 177 a 166 c 196 g 116 t 3 others

ORIGIN

Query Match 24.1%; Score 345.4; DB 14; Length 658;  
Best Local Similarity 86.2%; Pred. No. 3.5e-65;  
Matches 405; Conservative 0; Mismatches 59; Indels 6; Gaps 2;

QY 118 AACTGACCGGGAAGGTTGACAGGCGAG-GGAGAGAGGCCACAGACAGGCGACAG 176  
DB 194 ATCTAAGCGGGAAGGTTCTACAGGCGAGTGAAGGCGAGGAGCGCCAGGCCAG 253  
QY 177 GCAGGCTCCAGAGCGCGCGGATGAACTTCACGCGTGGGTTTCAAGCGCTCTAGGGGA 236  
DB 254 GCACGAGC-----ACCGAGGGGATGAACTTCACAGTGGGTTTCAAGCGCGCTAGGGGA 308  
QY 237 TGGCGACACATGACACACTTTGGAGAGCAGCTCATTTGCCCATCTCTCGAGATGTT 296  
DB 309 TGCACACAGCATGGACACACTGAGAGACAGCTCATCTGCCCATCTCGGAGATGTT 368  
QY 297 CTCACAGCCCGTGTGATCTTCTGCTGCAACACACACTGTGCGCAGTGTGCAACGA 356  
DB 369 CTCACACAGCATGGTGTATCTCTGCTGCAACACACACTGTGCGCAGTGTGCAACGA 428  
QY 357 CGTCTTCAGAGCCTCTAATCTCTGTGGCAATCCCGGGCTTCACACAGGTGTCTTCAAG 416  
DB 429 CGTCTTCAGAGCCTCTAATCTCTGTGGCAATCCCGGGCTTCACACAGGTGTCTTCAAG 488  
QY 417 AGGACGTTTCCGATGCCCATCTTGTAGGACAGAGTTTCTCTGACACAGCATGGTGTCA 476

DB 489 AGCGGTTTCCGCTGCCACATGTCAGGACGATAGAGTGTCTCTGACAGACAGGTCTCA 548  
QY 477 TGGCTTCAGCGGGAACCTGTCAGTGGAGACATCATTTACATCTACAGGAGGCTTC 536  
DB 549 CGGCTTCAGGGAACCTGTCAGTGGAGACATCATTTACATCTACAGGAGGCTTC 608  
QY 537 CCGGCACTGACCGCCAGGCTGGAACAGCAGCTCATCTGTGTGAGGAGCAG 586  
DB 609 CAGGCGCTGCACTTCACAGGCTGACAGACACTCATCTGTGAGGAGCATG 658

## RESULT 15

BI754358 793 bp mRNA linear EST 25-SEP-2001  
BI754358  
LOCUS 603026906F1 NIH\_MGC\_114 Homo sapiens cDNA clone IMAGE:5197560 5',  
DEFINITION mRNA sequence.

ACCESSION BI754358  
VERSION BI754358  
KEYWORDS GI:15745936  
SOURCE EST.  
human.

## ORGANISM

Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 793)  
NIH-MGC http://mgc.nci.nih.gov/  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cga@bbs-remail.nih.gov  
Tissue Procurement: Life Technologies, Inc.  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: L14M11494 row: k column: 01  
High quality sequence stop: 793.

## FEATURES

source

1. .793  
Location/Qualifiers

/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5197560"  
/clone\_lib="NIH\_MGC\_114"  
/lab\_host="DH10B"  
/note="Organ: brain; Vector: pCMV-Sport6; Site\_1: NotI;  
Site\_2: EcoRV (destroyed); RNA source anonymous pool of 6  
male brains; age range 23-27 yo. Library is oligo-dT  
primed and directionally cloned (EcoRV site is destroyed  
upon cloning). Average insert size 1.5 kb; insert size  
range 1-3 kb. Library is normalized and enriched for  
full-length clones and was constructed by C. Gruber  
(Invitrogen). Research Genetics tracking code 019. Note:  
this is a NIH\_MGC Library."

BASE COUNT 201 a 208 c 228 g 156 t

ORIGIN

Query Match 23.9%; Score 342.6; DB 13; Length 793;  
Best Local Similarity 72.1%; Pred. No. 1.6e-64;  
Matches 474; Conservative 0; Mismatches 179; Indels 4; Gaps 2;

QY 226 CTGCTAGGGATGCGCAACATGGAACACTTGGAGAGACAGTCAATTTGCCATGCG 285  
DB 140 CTGATCCAGGATGGGAATCCATGAGAACTTGGAGAGAGCTGATCTGCTATCTGC 199  
QY 286 CTGAGATGTTTTCACAGCCGCTGTGATCTTGGCTTCCCAACACAACTGTGCGGAG 345  
DB 200 CTGAGATGTTTTCACAGCCGCTGTGATCTTGGCTTCCCAACACAACTGTGCGGAG 259  
QY 346 TGTGCCAAGCAGCTCTTCAAGGCTCTAATCTCTGTGGCAATCCGGGGCTCCACAAG 405  
DB 260 TGTGCCAATGACATCTTCCAGGCTGTCAAAATCCCTACTTGACCAAGCCGGGGCTCA 319

QY 406 GTGCTCAGAGAGAGAGCTTTCCAGTCCCTGTAGAGACAGAGGTGCTGAGCAGG 465  
 || ||||| ||||| ||||| || || ||||| || |||||  
 Db 320 TCCATGCTCGAGAGCGCGTTTCCGCTGCGCCACAGAGTGCATGATCGT 379  
 || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 QY 466 CATGCTCTATGAGCTGACAGCGAAGCTGTAGTGAGACATTCATCTACAG 525  
 || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 380 CACGAGAGTACAGGCTGCAGAGAACCTGTGTGAGAACATCATCTACAAA 439  
 || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 QY 526 CAGAGCTCTCCGCGCAGCTGACAGCGCAAGGCTGAACACACTCTGTGTGAGAGC 565  
 || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 440 CAGAGCTCTCCAGTCCGCGCTGCAGAGAGGCA--GTCACCCATGTGCAAGGAGC 496  
 || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 QY 586 GAGAGCAGAAAGTACACATCTACTGCTGAGCTGAGAGTCCCACTGCTCTCTG 645  
 || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 497 GAAATGAGAAATCAACATCTACTGCTGAGCTGAGAGTCCCACTGCTCTCTG 556  
 || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 QY 646 AAGTTTCCGCGCGCAAGAGCTGTAGAGTGGCCCTCTGCGCCATTCACAAAGC 705  
 || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 557 AAGTGTGGAGTCCCAAGAGCTGTGAGAGTGGCCCAATTCAGAGTGTCTTCCAGG 616  
 || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 QY 706 CAGAGAGTGTGAGAGCTGAGAGTCCGATGCTGTGAGAGGCAATGACGCTGAG 765  
 || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 617 CAAAGAGTGAAGTGAATACATCTATCTGCTGAGTGTGAGAGGCAATGACGCTGAG 676  
 || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 QY 766 GCATGTATCACCAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAG 824  
 || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 677 ACCATCATCTACTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAG 736  
 || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 QY 825 GAACCACTGTAAACAGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAG 881  
 || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 737 CACGAGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAG 793  
 || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 16  
 LOCUS AM918753 378 bp mRNA linear EST 25-MAY-2000  
 DEFINITION Rattus norvegicus cdna clone R1EX57 5' end, mRNA sequence.  
 ACCESSION AM918753  
 VERSION AM918753.1 GI:8084538  
 KEYWORDS EST.  
 SOURCE Norway rat.  
 ORGANISM Rattus norvegicus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
 Rattus.  
 1 (bases 1 to 378)  
 Lee, N.H., Glodok, A., Chandra, I., Mason, T.M., Quackenbush, J.,  
 Kerlavage, A.R. and Adams, M.D.  
 Rat Genome Project: Generation of a Rat EST (RESE) Catalog & Rat  
 Gene Index  
 Unpublished (1998)  
 Contact: Lee, NH  
 The Institute for Genomic Research  
 9712, Medical Center Drive, Rockville, MD 20850, USA  
 Tel: (301)-838-3529  
 Fax: (301)-838-0208  
 Email: nhlee@tigr.org  
 This clone is available through the ATCC, contact the ATCC  
 tel#703-365-2700 for further information  
 Seq primer: M13 Reverse.  
 Location/Qualifiers  
 1..378  
 /organism="Rattus norvegicus"  
 /db\_xref="taxon:10116"  
 /clone.lib="RG1EX57"  
 /clone.lib="Rat gene index, normalized rat, norvegicus,  
 Bonto Soares"  
 /tissue\_type="mix - brain, ovary, placenta, kidney, lung,  
 liver, embryo, heart, muscle, spleen"  
 /lab\_host="SOLR"  
 /note="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2:  
 XhoI; Estimated insert size approx.1 kb"

BASE COUNT 87 a 100 c 122 g 69 t  
 ORIGIN  
 Query Match 23.9%; Score 341.8; DB 10; Length 378;  
 Best Local Similarity 94.2%; Pred. No. 1.7e-64;  
 Matches 355; Conservative 0; Mismatches 22; Indels 0; Gaps 0;  
 QY 599 TCACATCTACTGCTGAGCTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAG 658  
 || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 2 TTACATCTACTGCTGAGCTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAG 61  
 || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 QY 659 CCCAAGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAG 718  
 || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 62 CCCAAGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAG 121  
 || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 QY 719 TGAGCATGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAG 778  
 || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 122 TGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAG 181  
 || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 QY 779 AGATGAGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAG 838  
 || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 182 AGATGAGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAG 241  
 || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 QY 839 ACCAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAG 898  
 || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 242 ACCAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAG 301  
 || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 QY 899 CACTGCGCGGAG 958  
 || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 302 CGCTGCTCGGAG 361  
 || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 QY 959 GAGACCACTGTGAGAGG 975  
 || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 362 GAGACCACTGTGAGAGG 378  
 || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 17  
 LOCUS BG375657/c 440 bp mRNA linear EST 12-MAR-2001  
 DEFINITION UI-R-CSO-btd-h-03-0-UI-s1 UI-R-CSO Rattus norvegicus cdna clone  
 UI-R-CSO-btd-h-03-0-UI 3', mRNA sequence.  
 ACCESSION BG375657  
 VERSION BG375657.1 GI:13300129  
 KEYWORDS EST.  
 SOURCE Norway rat.  
 ORGANISM Rattus norvegicus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
 Rattus.  
 1 (bases 1 to 440)  
 Bonaldo, M.F., Lennon, G. and Soares, M.B.  
 Normalization and subtraction: two approaches to facilitate gene  
 discovery  
 Genome Res. 6 (9), 791-806 (1996)  
 9704477  
 Contact: Soares, MB  
 Program for Rat Gene Discovery and Mapping  
 University of Iowa  
 451 Eckstein Medical Research Building Iowa City, IA 52242, USA  
 Tel: 319 335 8250  
 Fax: 319 335 9565  
 Email: msoares@blue.weeg.uiowa.edu  
 The sequence contained an oligo-dt track that was present in the  
 oligonucleotide that was used to prime the synthesis of first  
 strand cdna and therefore this may represent a bonafide poly A  
 tail. The sequence tag present in the cdna between the NotI site  
 and the oligo-dt track served to verify it as a clone from the  
 normalized rat heart pool library cdna library Preparation: M.B.  
 Soares Lab Clone distribution: clones will be available through  
 Research Genetics (www.resgen.com)  
 Seq primer: M13 Forward  
 POLYA=yes.  
 Location/Qualifiers

## FEATURES

FEATURES  
 source  
 1..378  
 /organism="Rattus norvegicus"  
 /db\_xref="taxon:10116"  
 /clone.lib="RG1EX57"  
 /clone.lib="Rat gene index, normalized rat, norvegicus,  
 Bonto Soares"  
 /tissue\_type="mix - brain, ovary, placenta, kidney, lung,  
 liver, embryo, heart, muscle, spleen"  
 /lab\_host="SOLR"  
 /note="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2:  
 XhoI; Estimated insert size approx.1 kb"



source

1. .440  
/organism="Rattus norvegicus"  
/strain="Sprague-Dawley"  
/db\_xref="taxon:10116"  
/clone="UI-R-CS0-btd-b-03-0-UI"  
/clone\_11b="UI-R-CS0"  
/dev\_stage="ADULT"  
/lab\_host="DHI0B (Life Technologies)"  
/note="Vector: p7T3D-Pac (Pharmacia) with a modified polylinker. Site\_1: Not I; Site\_2: Eco RI; The UI-R-CS0 library is a normalized library constructed from the following rat heart tissues: embryonic day 17, embryonic day 19, embryonic day 21, adult day 1, adult day 12, adult day 75, adult day 200. For a detailed description of the library from which this clone was derived, please visit our web site at [rattest.eng.uiowa.edu](http://rattest.eng.uiowa.edu). The subtraction has been previously described in (Bonaldo, Lennon and Soares, Genome Research 6:791-806, 1996)  
TAG\_Lib=UI-R-CS0  
TAG\_Tissue=rat heart pool  
TAG\_SEQ=ATAGATAC"

BASE COUNT 69 a 144 c 113 g 114 t

ORIGIN

Query Match 23.9%; Score 341.4; DB 12; Length 440;  
Best Local Similarity 90.3%; Pred. No. 2.2e-64;  
Matches 391; Conservative 0; Mismatches 31; Indels 11; Gaps 2;

OY 999 CATCCAGTCCATGAGAGAGCCGCGATGCTCTCTACCTCCAGCAGCAAGAGCTGAT 1058  
|||||  
DB 440 CATCCAGTCCATGAGAGAGCCGCGATGCTCTCTACCTCCAGCAGCAGAGCTGAT 381  
|||||

OY 1059 CAACAAGTCCGAGGAGGATGCAAGTGGAGTGGAGAGAGCCGAGCAGGCTATGA 1118  
|||||  
DB 380 CAACAAGTCCGAGGAGGATGCAAGTGGAGTGGAGAGCCGAGCAGGCTATGA 321  
|||||

OY 1119 GAGCATGAGCAATTCCTGTGAGCTGGAGCAGTGGCCGAATGTTGGCAACCATCGA 1178  
|||||  
DB 320 GAGCATGAGCAATTCCTGTGAGCTGGAGCAGTGGCCGAATGTTGGCAACCATCGA 261  
|||||

OY 1179 CTTCAGAGCCGCGCGCTGGGGATGAGAGATGACGACATCTTGGATGGGAGGA 1238  
|||||  
DB 260 CTTCAGAGCCGCGCGCTGGGGATGAGAGATGACGACATCTTGGATGGGAGGA 201  
|||||

OY 1239 GGGCATGGGGCTGGAGAGAGGCGCTGACGTGCCAAGGCTCAGGCTGACCTG 1298  
|||||  
DB 200 GGGCATGGGGCTGGAGAGAGGCGCTGACGTGCCAAGGCTCAGGCTGACCTG 147  
|||||

OY 1299 ACCGCACTGTATCCAGAGCGCACACCAGGAGGAGGAGGATGCTGAGATCTGC 1358  
|||||  
DB 146 ATCCGCACTGTATCCAGAGCGCACACCAGGAGGAGGATGCTGAGATCTGC 92  
|||||

OY 1359 GCAGAGACCAAGCGCCAGCAGAGCTGGCTTCCGCGCCCGGGAGGTTCTCAATAAGG 1418  
|||||  
DB 91 GCAGAGACCAAGCGCCAGCAGAGCTGGCTTCCGCGCCCGGGAGGTTCTCAATAAGG 32  
|||||

OY 1419 ACTCAGTGTGCC 1431  
|||||  
DB 31 ACTCAGTGTGTGCC 19  
|||||

RESULT 18  
BG379943/C 441 bp mRNA linear EST 12-MAR-2001  
DEFINITION UI-R-CS0-btd-b-01-0-UI.s1 UI-R-CS0 Rattus norvegicus cDNA clone.  
ACCESSION BG379943  
VERSION BG379943.1 GI:13304415  
KEYWORDS EST.  
SOURCE Norway rat.  
ORGANISM Rattus norvegicus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murine;

REFERENCE 1 (bases 1 to 441)  
AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.  
TITLE Normalization and subtraction: two approaches to facilitate gene discovery  
JOURNAL Genome Res. 6 (9), 791-806 (1996)  
MEDLINE 97044477  
COMMENT Contact: Soares, MB  
Program for Rat Gene Discovery and Mapping  
University of Iowa  
451 Eckstein Medical Research Building Iowa City, IA 52242, USA  
Tel: 319 335 8250  
Fax: 319 335 9565  
Email: [masoares@blue.weeg.uiowa.edu](mailto:masoares@blue.weeg.uiowa.edu)  
The sequence contained an oligo-dT track that was present in the oligonucleotide that was used to prime the synthesis of first strand cDNA and therefore this may represent a bonafide poly A tail. cDNA library preparation: M.B. Soares Lab Clone distribution: clones will be available through Research Genetics ([www.resgen.com](http://www.resgen.com))  
Seq primer: M13 Forward  
POLYA=yes

FEATURES

source

1. .441  
/organism="Rattus norvegicus"  
/strain="Sprague-Dawley"  
/db\_xref="taxon:10116"  
/clone="UI-R-CS0-btd-b-01-0-UI"  
/clone\_11b="UI-R-CS0"  
/dev\_stage="ADULT"  
/lab\_host="DHI0B (Life Technologies)"  
/note="Vector: p7T3D-Pac (Pharmacia) with a modified polylinker. Site\_1: Not I; Site\_2: Eco RI; The UI-R-CS0 library is a normalized library constructed from the following rat heart tissues: embryonic day 17, embryonic day 19, embryonic day 21, adult day 1, adult day 12, adult day 75, adult day 200. For a detailed description of the library from which this clone was derived, please visit our web site at [rattest.eng.uiowa.edu](http://rattest.eng.uiowa.edu). The subtraction has been previously described in (Bonaldo, Lennon and Soares, Genome Research 6:791-806, 1996)  
TAG\_SEQ=None found"

BASE COUNT 69 a 148 c 110 g 114 t

ORIGIN

Query Match 23.8%; Score 340.6; DB 12; Length 441;  
Best Local Similarity 90.7%; Pred. No. 3.3e-64;  
Matches 389; Conservative 0; Mismatches 29; Indels 11; Gaps 2;

OY 1003 CATCCATGAGAGAGCCGCGATGCTCTCTACCTCCAGCAGCAGGCTGATCAAC 1062  
|||||  
DB 441 CATCCATGAGAGAGCCGCGATGCTCTCTACCTCCAGCAGCAGGCTGATCAAC 382  
|||||

OY 1063 AAGTTCGGGGCAATGTCGAAGTGGAGCTGGAGAGCCGCGAGGCTATGAGAGC 1122  
|||||  
DB 381 AAGTTCGGGGCAATGTCGAAGTGGAGCTGGAGAGCCGCGAGGCTATGAGAGC 322  
|||||

OY 1123 ATGAGCAATTCCTGTGAGCGTGGAGACAGTGGCCGAATGTTGGCAACCATGACTTC 1182  
|||||  
DB 321 ATGAGCAATTCCTGTGAGCGTGGAGACAGTGGCCGAATGTTGGCAACCATGACTTC 262  
|||||

OY 1183 CAGCCGGCGCGCTGGGGATGAGAGATGACGACATGCTTGGATGGGAGAGGCG 1242  
|||||  
DB 261 CAGCCGGCGCGCTGGGGATGAGAGATGACGACATGCTTGGATGGGAGAGGCG 202  
|||||

OY 1243 AATGCGGGCTGGAGAGAGCGGCTGGAGAGTCCAGAGGCTCAGGCTGCTGAGACC 1302  
|||||  
DB 201 AACACAGGCTGGAGAGAGCGGCTGGAGAGTCCAGAGGCTCAGGCTGCTGAGACC 148  
|||||

OY 1303 GACTTGATCCAGAGCGCACACCCGAGGAGCCAGGAGATGCTGAGATCTGCGCAG 1362  
|||||  
DB 147 GACTTGATCCAGAGCGCACACCCGAGGAGCCAGGAGATGCTGAGATCTGCGCAG 93  
|||||

OY 1363 AGACCAAGCGCGCCAGCAGAGCTGGCTTCCGCGCCCGGGAGAGGTTCTCAATAAGGACTC 1422  
|||||

Db 92 AGACCATCGCGCTACCATCTCGCGCTCCACCCGAGGGATGTTCCATTAAGACTC 33  
 QY 1423 AAGTGTCCC 1431  
 Db 32 AAGTGTCCC 24

RESULT 19  
 BG764060 868 bp mRNA linear EST 15-MAY-2001  
 BG764060  
 LOCUS 60273706991 NIH\_MGC\_49 Homo sapiens cDNA clone IMAGE:4862321 5',  
 mRNA sequence.  
 BG764060  
 ACCESSION BG764060.1 GI:14074713  
 VERSION  
 KEYWORDS  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 868)  
 AUTHORS NIH-MGC http://mgc.nci.nih.gov/.  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cgapbs-remail.nih.gov  
 Tissue Procurement: ATCC/DC/DTP  
 CDNA Library Preparation: Ling Hong/Rubin Laboratory  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov  
 Plate: LICM1722 row: j column: 18  
 High quality sequence stop: 863.

FEATURES  
 source  
 1..868  
 location/Qualifiers  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:4862321"  
 /clone\_1lb="NIH\_MGC\_49"  
 /tissue\_type="melanotic melanoma, high MDR (cell line)"  
 /lab\_host="DH10B (phage-resistant)"  
 /note="Organ: skin; Vector: pORF7; Site: 1: XhoI; Site: 2:  
 EcoRI; CDNA made by oligo-dt priming. Directionally cloned  
 into EcoRI/XhoI sites using the following 5' adaptor:  
 GGCACGAG(G). Size selected >500bp for average insert size  
 1.8kb. Library constructed by Ling Hong in the Laboratory  
 of Gerald M. Rubin (University of California, Berkeley)  
 using ZAP-cDNA synthesis kit (Stratagene) and Superscript  
 II RT (Life Technologies). Note: this is a NIH\_MGC  
 Library."

BASE COUNT 224 a 222 c 252 g 170 t  
 ORIGIN

Query Match 23.38; Score 333; DB 12; Length 868;  
 Best Local Similarity 70.3%; Pred. No. 2; Le-62;  
 Matches 517; Conservative 0; Mismatches 210; Indels 8; Gaps 5;

QY 226 CTGTAGGAGTGGCGACATGAGACACTTGAGAGAGCGCTCATTTGCCCATCTGC 285  
 Db 138 CTGATCCAGATGGGAATCCCATGAGACTTGAGAGAGCGCTCATCTGCCTATCTGC 197  
 QY 286 CTGAGATGTTTCCAGACCCGCTGTATCTTGGCTTCCCAACACACCTGTGCGCAAG 345  
 Db 198 CTGAGATGTTTACAGACAGAGTGTCTTGGCGTCCAGCAACACTGTGCGGGAAG 257  
 QY 346 TGTGCCACAGACGCTTCCAGAGCTTAACTCTGTGCGCAATCCCGGGGCTCCACAAG 405  
 Db 258 TGTGCAATGACATCTTCCAGGCTGCAAAATCCCTACTGACCAAGCGGGGAGCTCAGTG 317  
 QY 406 GTGTCTTCCAGAGACCTTCCGATGCTGTAGGCGAGAGGTGTCTCTGACAGG 465

Db 318 TCATGTCGTGAGAGCCCTTTCCGCTGCCCCACCTGCCCCACGAGGTATGATGATCT 377  
 QY 466 CATGGTGTCTATGGCCCTCAGCGGAACCTGTATGAGAGAACATATTGACATTTCAAG 525  
 Db 378 CAGGAGAGTGTACGGCTCAGAGGAACCTGTGTGGAGAACATATGACATTTCAAAA 437  
 QY 526 CAGGAGTCTCCCGCCCACTGACAGCCCAAGGCTGAACAGCACTATGTGTAGAGAC 585  
 Db 438 CAGGAGTCTCCAGTGGCGCCCTGCGAAG--GGCAGTCAACCCATGTGCAAGAGCAC 494  
 QY 586 GAGGACGAGAGAGATCA--CATCTACTGCTGAGTGCAGAGTGGCGCACCTGCTCTGTG 644  
 Db 495 GAAGATGAGAAATCAACACATCTACTGTCTACGTGTGAGTGGCCCACTGCTCATGTG 554  
 QY 645 CAAGTTTTCGGCCGCCCAAGAGACTGTGAGTGGCCCTGCTGCCACCATTTTACAAAG 704  
 Db 555 CAAGTGTGTTGGGATTCACAAGGCTGCGAGGTGCCCCATGTGACAGAGTCTTCCAGG 614  
 QY 705 CCAGAGAGTGTGAGTGTGAGCGATGCGATGCGTGTGGCGGCGCATATGACCGTGTGA 764  
 Db 615 ACAGAAAGCTAATGATGATTAATCTATCTCATGTGTGTGGGGAATGACCGTGTGA 674  
 QY 765 GGCAGTATCACCCAGATGAGAGAGTGTGCCAGACATTGAGACAAACAGCCGAGACA 824  
 Db 675 GACCATCATCTACCTCAGCTGAGAGATTCCTCCGTGAGTGAACCAAGAGAACAGTACAGGT 734  
 QY 825 GAAGCACTGTTAATCAG--AGTTTGCAGACCTGTGCGCGGTTTGTGAGAGCGCAAG 882  
 Db 735 AAAGAAAGAGCTGAGCCAGACAGTGTGACAGTTAGTATGTCATCTGATGAGAAAGA 794  
 QY 883 GGCACACTGCTTCAAGACACTGCGCGGAGCAGAGGAGAGTTCACACCGCTGG--GGG 941  
 Db 795 AGTATGTTGTGTGCGAGGA--TCACGAGAGAGAGAGACAAAGCTTAATGATCCGAGG 853  
 QY 942 CTTATCCGCGCAGTA 956  
 Db 854 CTTATCCGCGCAGTA 868

RESULT 20  
 BM488618  
 LOCUS 495 bp mRNA linear EST 07-FEB-2002  
 DEFINITION pgm2n.pk008.d8 Normalized Chicken Breast Muscle, Leg Muscle, and  
 Epiphyseal Growth Plate CDNA Library (pgm2n) Gallus gallus cDNA  
 clone pgm2n.pk008.d8 5', similar to ref|NP\_067422.1| (NM\_021447)  
 muscle-specific RING-finger protein; 4930486E09RX [Mus musculus]  
 gb|AA03076.1| (AF294790) RING-finger protein MDRF [Mus musculus],  
 mRNA sequence.  
 BM488618  
 ACCESSION BM488618.1 GI:18609549  
 VERSION  
 KEYWORDS  
 SOURCE chicken.  
 ORGANISM Gallus gallus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Archosauria; Aves; Neognathae; Galliformes; Phasianidae;  
 Phasianinae; Gallus.  
 REFERENCE 1 (bases 1 to 495)  
 AUTHORS Cogburn, L.A. and Monsonego-Ornan, E.  
 ESTs from Normalized Chicken Breast Muscle, Leg Muscle, and  
 Epiphyseal Growth Plate CDNA Library, USDA/IRAFs Animal Genome  
 Project  
 JOURNAL Unpublished (2002)  
 COMMENT Contact: Larry A. Cogburn  
 University of Delaware  
 Townsend Hall, Newark, DE 19717, USA  
 Tel: 302-831-1335  
 Fax: 302-831-2822  
 Email: cogburnudel.edu, www.chickest.udel.edu.

FEATURES  
 source  
 1..495  
 location/Qualifiers  
 /organism="Gallus gallus"  
 /strain="Commercial broiler and Ottawa Res. Centre  
 Strains 90 & 21"





GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 3, 2002, 09:13:42 ; Search time 59 Seconds  
(without alignments) updates/sec  
7438.211 Million cell

Title: US-09-908-988B-1

Perfect score: 1431  
Sequence: 1 aagagagtgtacacagagtgt.....ataaagactcaagtgtccc 1431Scoring table: OLIGO-MUC  
Gapop 60.0 , Gapext 60.0

Searched: 441362 seqs, 15338381 residues

Word size: 15

Total number of hits satisfying chosen parameters: 744

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database:

Issued Patents, NA: \*

- 1: /cgn2\_6/ptodata/1/ina/5A.COMB.seq:\*
- 2: /cgn2\_6/ptodata/1/ina/5B.COMB.seq:\*
- 3: /cgn2\_6/ptodata/1/ina/6A.COMB.seq:\*
- 4: /cgn2\_6/ptodata/1/ina/6B.COMB.seq:\*
- 5: /cgn2\_6/ptodata/1/ina/PCRTUS.COMB.seq:\*
- 6: /cgn2\_6/ptodata/1/ina/Backfile1.seq:\*

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	17	1.2	39	3	US-08-767-942A-36
2	17	1.2	102	1	US-08-047-041A-15
3	17	1.2	133	1	US-08-047-041A-2
4	17	1.2	133	2	US-08-795-006A-22
5	17	1.2	133	4	US-09-184-073-22
6	17	1.2	320	3	US-09-157-177-116
7	17	1.2	716	4	US-08-998-416-1129
8	17	1.2	732	4	US-08-998-416-927
9	17	1.2	800	1	US-08-472-239-1
10	17	1.2	1070	4	US-09-414-436-4
11	17	1.2	1215	1	US-08-347-792-20
12	17	1.2	1215	1	US-08-431-357-20
13	17	1.2	1215	4	US-08-392-542-26
14	17	1.2	1215	4	US-08-894-327-26
15	17	1.2	1215	4	US-09-685-027-26
16	17	1.2	1215	5	PCT-US95-15353-20
17	17	1.2	1303	1	US-08-047-041A-13
18	17	1.2	1307	1	US-08-047-041A-12
19	17	1.2	1317	1	US-07-912-011-1
20	17	1.2	1317	1	US-08-347-792-1
21	17	1.2	1317	1	US-08-431-357-1
22	17	1.2	1317	2	US-08-697-221-1
23	17	1.2	1317	4	US-08-392-542-1
24	17	1.2	1317	4	US-08-894-327-1
25	17	1.2	1317	4	US-09-147-751-8
26	17	1.2	1317	4	US-09-305-914-1
27	17	1.2	1317	4	US-09-685-027-1

28	17	1.2	1317	5	PCT-US95-15353-1	Sequence 1, Appli
29	17	1.2	1683	1	US-07-945-283-3	Sequence 3, Appli
30	17	1.2	1710	1	US-07-903-103-3	Sequence 3, Appli
31	17	1.2	1710	1	US-08-044-619A-3	Sequence 3, Appli
32	17	1.2	1710	1	US-08-283-911-3	Sequence 3, Appli
33	17	1.2	1710	1	US-08-245-500A-4	Sequence 4, Appli
34	17	1.2	1710	1	US-08-390-546-4	Sequence 4, Appli
35	17	1.2	1710	1	US-08-390-479A-4	Sequence 4, Appli
36	17	1.2	1710	1	US-08-557-393-4	Sequence 4, Appli
37	17	1.2	1710	1	US-08-390-516C-4	Sequence 4, Appli
38	17	1.2	1710	1	US-08-390-517A-4	Sequence 4, Appli
39	17	1.2	1710	1	US-08-390-515A-4	Sequence 4, Appli
40	17	1.2	1710	2	US-08-801-718-4	Sequence 4, Appli
41	17	1.2	1710	3	US-09-073-567-12	Sequence 12, Appli
42	17	1.2	1710	4	US-09-170-159A-4	Sequence 4, Appli
43	17	1.2	1710	4	US-09-480-718-45	Sequence 45, Appli
44	17	1.2	1722	4	US-09-385-028-15	Sequence 15, Appli
45	17	1.2	1757	3	US-08-094-071-1	Sequence 1, Appli

## ALIGNMENTS

RESULT 1  
US-08-767-942A-36  
Sequence 36, Application US/08767942A  
Patent No. 6068982

GENERAL INFORMATION:  
APPLICANT: Rolfe, Mark  
APPLICANT: Chin, M. Isabel  
APPLICANT: Berlin, Vivian  
APPLICANT: Damagnez, Veronique  
APPLICANT: Draetta, Giulio  
APPLICANT: Guillaume, Cottarel  
TITLE OF INVENTION: UNROUTIN COMUGATING ENZYMES  
NUMBER OF SEQUENCES: 45  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: FOLEY, HOAG & ELIOT LLP  
STREET: One Post Office Square  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109-2170

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/767, 942A  
FILING DATE: 17-DEC-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Vincent, Matthew P.  
REGISTRATION NUMBER: 36,709  
REFERENCE/DOCKET NUMBER: MIV-029.04  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-832-7000  
TELEFAX: 617-832-1000

INFORMATION FOR SEQ. ID NO: 36:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 39 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: other nucleic acid  
DESCRIPTION: /desc = "primer"

US-08-767-942A-36

Query Match 1.28; Score 17; DB 3; Length 39;  
Best local Similarity 100.0%; Pred. No. 87;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1007 CCATGAGGAGCCGACG 1023

Db 14 CCATGAGAGAGCCGAG 30

|||||

## RESULT 2

US-08-047-041A-15

Sequence 15, Application US/08047041A

Patent No. 5527676

GENERAL INFORMATION:

APPLICANT: Vogelstein, Bert

APPLICANT: Baker, Suzanne J.

APPLICANT: Fearon, Eric R.

APPLICANT: Nigro, Janice M.

TITLE OF INVENTION: Detection of Loss of the Wild-Type p53

TITLE OF INVENTION: Gene

NUMBER OF SEQUENCES: 28

CORRESPONDENCE ADDRESSES:

ADDRESSEE: Banner & Allegretti, Ltd.

STREET: 1001 G Street, N.W.

CITY: Washington

STATE: D.C.

COUNTRY: USA

ZIP: 20001-4597

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/047,041A

FILING DATE: 22-MAR-1993

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/928,661

FILING DATE: 17-AUG-1992

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/446,584

FILING DATE: 06-DEC-1989

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/330,566

FILING DATE: 29-MAR-1989

ATTORNEY/AGENT INFORMATION:

NAME: Kagan, Sarah A.

REGISTRATION NUMBER: 32,141

REFERENCE/DOCKET NUMBER: 01107,42917

TELECOMMUNICATION INFORMATION:

TELEPHONE: 202-508-9100

TELEFAX: 202-508-9299

INFORMATION FOR SEQ ID NO: 15:

SEQUENCE CHARACTERISTICS:

LENGTH: 102 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: cDNA

HYPOHETICAL: NO

ANTI-SENSE: NO

ORIGINAL SOURCE:

ORGANISM: Homo sapiens

POSITION IN GENOME:

CHROMOSOME/SEGMENT: exon 2

PUBLICATION INFORMATION:

AUTHORS: Lamb,

JOURNAL: Mol. Cell. Biol.

VOLUME: 6

ISSUE: 5

PAGES: 1379-1385

DATE: 1986

US-08-047-041A-15

Query Match

Best Local Similarity 1.2%; Score 17; DB 1; Length 102;

Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1007 CCATGAGAGAGCCGAG 1023

|||||

Db 27 CCATGAGAGAGCCGAG 43

## RESULT 3

US-08-047-041A-2

Sequence 2, Application US/08047041A

Patent No. 5527676

GENERAL INFORMATION:

APPLICANT: Vogelstein, Bert

APPLICANT: Baker, Suzanne J.

APPLICANT: Fearon, Eric R.

APPLICANT: Nigro, Janice M.

TITLE OF INVENTION: Detection of Loss of the Wild-Type p53

TITLE OF INVENTION: Gene

NUMBER OF SEQUENCES: 28

CORRESPONDENCE ADDRESSES:

ADDRESSEE: Banner & Allegretti, Ltd.

STREET: 1001 G Street, N.W.

CITY: Washington

STATE: D.C.

COUNTRY: USA

ZIP: 20001-4597

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/047,041A

FILING DATE: 22-MAR-1993

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/928,661

FILING DATE: 17-AUG-1992

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/446,584

FILING DATE: 06-DEC-1989

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/330,566

FILING DATE: 29-MAR-1989

ATTORNEY/AGENT INFORMATION:

NAME: Kagan, Sarah A.

REGISTRATION NUMBER: 32,141

REFERENCE/DOCKET NUMBER: 01107,42917

TELECOMMUNICATION INFORMATION:

TELEPHONE: 202-508-9100

TELEFAX: 202-508-9299

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 133 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: cDNA

ORIGINAL SOURCE:

ORGANISM: Homo sapiens

POSITION IN GENOME:

CHROMOSOME/SEGMENT: exon 2

PUBLICATION INFORMATION:

AUTHORS: Buchman, V. L.

TITLE: A variation in the structure of the

JOURNAL: Gene

VOLUME: 70

PAGES: 245-252

DATE: 1988

US-08-047-041A-2

Query Match

Best Local Similarity 1.2%; Score 17; DB 1; Length 133;

Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1007 CCATGAGAGAGCCGAG 1023  
|||||  
DB 42 CCATGAGAGAGCCGAG 58

## RESULT 4

US-08-795-006A-22  
Sequence 22, Application US/08795006A  
Patent No. 5840579  
GENERAL INFORMATION:  
APPLICANT: Boeke, Jef  
APPLICANT: Brachmann, Rainer  
TITLE OF INVENTION: NUCLEIC ACIDS ENCODING P53  
TITLE OF INVENTION: MUTATIONS WHICH SUPPRESS P53 CANCER MUTATIONS  
NUMBER OF SEQUENCES: 32  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Banner & Witcoff  
STREET: 1001 G Street, NW  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20001  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/795,006A  
FILING DATE: 05-FEB-1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
ATTORNEY/AGENT INFORMATION:  
NAME: Kagan, Sarah A  
REGISTRATION NUMBER: 32141  
REFERENCE/DOCKET NUMBER: 01107.03170  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-508-9100  
TELEFAX: 202-508-9299  
TELEX:  
INFORMATION FOR SEQ ID NO: 22:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 133 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-795-006A-22

Query Match 1.2%; Score 17; DB 2; Length 133;  
Best Local Similarity 100.0%; Pred. No. 85;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1007 CCATGAGAGAGCCGAG 1023  
|||||  
DB 42 CCATGAGAGAGCCGAG 58

## RESULT 5

US-09-184-073-22  
Sequence 22, Application US/09184073  
Patent No. 6183964  
GENERAL INFORMATION:  
APPLICANT: Boeke, Jef  
APPLICANT: Brachmann, Rainer  
TITLE OF INVENTION: NUCLEIC ACIDS ENCODING P53  
TITLE OF INVENTION: MUTATIONS WHICH SUPPRESS P53 CANCER MUTATIONS  
NUMBER OF SEQUENCES: 32  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Banner & Witcoff

STREET: 1001 G Street, NW  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20001  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/184,073  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/795,006  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Kagan, Sarah A  
REGISTRATION NUMBER: 32141  
REFERENCE/DOCKET NUMBER: 01107.03170  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-508-9100  
TELEFAX: 202-508-9299  
TELEX:  
INFORMATION FOR SEQ ID NO: 22:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 133 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-09-184-073-22

Query Match 1.2%; Score 17; DB 4; Length 133;  
Best Local Similarity 100.0%; Pred. No. 85;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1007 CCATGAGAGAGCCGAG 1023  
|||||  
DB 42 CCATGAGAGAGCCGAG 58

## RESULT 6

US-09-157-177-116  
Sequence 116, Application US/09157177  
Patent No. 6090558  
GENERAL INFORMATION:  
APPLICANT: Butler, John M.  
APPLICANT: Li, Jia  
APPLICANT: Monforte, Joseph A.  
APPLICANT: Becker, Christopher H.  
TITLE OF INVENTION: DNA TYPING BY MASS SPECTROMETRY WITH POLYMORPHIC DNA  
FILE REFERENCE: GETR:017/GEPR017P  
CURRENT APPLICATION NUMBER: US/09/157,177  
CURRENT FILING DATE: 1998-09-18  
EARLIER APPLICATION NUMBER: 60/059,415  
EARLIER FILING DATE: 1997-09-19  
NUMBER OF SEQ ID NOS: 135  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 116  
LENGTH: 320  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-157-177-116

Query Match 1.2%; Score 17; DB 3; Length 320;  
Best Local Similarity 100.0%; Pred. No. 83;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 27 ATAGACAGGGGTGAGAG 43  
|||||  
DB 237 ATAGACAGGGGTGAGAG 253

RESULT 7  
US-08-998-416-1129  
; Sequence 1129, Application US/08998416  
; Patent No. 6239264  
; GENERAL INFORMATION:  
; APPLICANT: Philippesen, Peter  
; APPLICANT: Pohlmann, Rainer  
; APPLICANT: Steiner, Sabine  
; APPLICANT: Mohr, Christine  
; APPLICANT: Wendland, Jurgen  
; APPLICANT: Knechtle, Philipp  
; APPLICANT: Reibischung, Corinne  
; TITLE OF INVENTION: GENOMIC DNA SEQUENCES OF ASHBYA GOSYPPII  
; TITLE OF INVENTION: AND USES THEREOF  
; NUMBER OF SEQUENCES: 1152  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: No. 6239264artis Corporation  
; STREET: 3054 Cornwalis Road  
; CITY: Research Triangle Park  
; STATE: No. 6239264th Carolina  
; COUNTRY: USA  
; ZIP: 27709  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/998,416  
; FILING DATE: 24-DEC-1997  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: CH 0016/97  
; FILING DATE: 31-DEC-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Meigs, J. Timothy  
; REGISTRATION NUMBER: 38,241  
; REFERENCE/DOCKET NUMBER: PF/5-30306/A/CGC1976  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 919-541-8587  
; TELEFAX: 919-541-8689  
; INFORMATION FOR SEQ ID NO: 1129:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 716 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
; ORIGINAL SOURCE:  
; ORGANISM: PAG1688RP  
; US-08-998-416-1129

Query Match 1.2%; Score 17; DB 4; Length 716;  
Best Local Similarity 100.0%; Pred. No. 82;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1192 GCCGCTGGGGATGAGA 1208  
|||||

DB 526 GCCGCTGGGGATGAGA 542

RESULT 8  
US-08-998-416-927  
; Sequence 927, Application US/08998416  
; Patent No. 6239264  
; GENERAL INFORMATION:  
; APPLICANT: Philippesen, Peter  
; APPLICANT: Pohlmann, Rainer  
; APPLICANT: Steiner, Sabine  
; APPLICANT: Mohr, Christine  
; APPLICANT: Wendland, Jurgen

APPLICANT: Knechtle, Philipp  
APPLICANT: Reibischung, Corinne  
TITLE OF INVENTION: GENOMIC DNA SEQUENCES OF ASHBYA GOSYPPII  
TITLE OF INVENTION: AND USES THEREOF  
NUMBER OF SEQUENCES: 1152  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: No. 6239264artis Corporation  
STREET: 3054 Cornwalis Road  
CITY: Research Triangle Park  
STATE: No. 6239264th Carolina  
COUNTRY: USA  
ZIP: 27709  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/998,416  
FILING DATE: 24-DEC-1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: CH 0016/97  
FILING DATE: 31-DEC-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Meigs, J. Timothy  
REGISTRATION NUMBER: 38,241  
REFERENCE/DOCKET NUMBER: PF/5-30306/A/CGC1976  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 919-541-8587  
TELEFAX: 919-541-8689  
INFORMATION FOR SEQ ID NO: 927:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 732 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
ORIGINAL SOURCE:  
ORGANISM: PAG1576RP  
US-08-998-416-927

Query Match 1.2%; Score 17; DB 4; Length 732;  
Best Local Similarity 100.0%; Pred. No. 82;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1192 GCCGCTGGGGATGAGA 1208  
|||||

DB 526 GCCGCTGGGGATGAGA 542

RESULT 9  
US-08-472-239-1  
; Sequence 1, Application US/08472239  
; Patent No. 5728526  
; GENERAL INFORMATION:  
; APPLICANT: GEORGE, Jr., Albert L.  
; APPLICANT: BHATNAGAR, Satish K.  
; APPLICANT: NAZARENKO, Irena  
; TITLE OF INVENTION: METHOD FOR ANALYZING A NUCLEOTIDE  
; TITLE OF INVENTION: SEQUENCE AND KIT THEREFOR  
; NUMBER OF SEQUENCES: 2  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS  
; STREET: P.O. Box 1404  
; CITY: Alexandria  
; STATE: Virginia  
; COUNTRY: United States  
; ZIP: 22313-1404  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: PC-DOS/MS-DOS



```
; SOFTWARE: Patentln Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/472.239
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Mukai, Robert G.
; REGISTRATION NUMBER: 28,531
; REFERENCE/DOCKET NUMBER: 020160-229
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6620
; TELEFAX: (703) 836-2021
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 800 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-472-239-1

Query Match          1.2%; Score 17; DB 1; Length 800;
Best Local Similarity 100.0%; Pred. No. 82;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1007 CCATGGAGGAGCCGCAG 1023
DB 115 CCATGGAGGAGCCGCAG 131

RESULT 10
US-09-414-436-4
; Sequence 4, Application US/09414436
; Patent No. 6294384
; GENERAL INFORMATION:
; APPLICANT: Dell'Acqua, Giorgio
; APPLICANT: Mann, Michael J.
; APPLICANT: Dzaou, Victor J.
; TITLE OF INVENTION: Compositions and Methods Based Upon an Isoform of p53
; FILE REFERENCE: P53F
; CURRENT APPLICATION NUMBER: US/09/414,436
; CURRENT FILING DATE: 1999-10-07
; EARLIER APPLICATION NUMBER: 60/103,849
; EARLIER FILING DATE: 1999-10-09
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 4
; LENGTH: 1070
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-414-436-4

Query Match          1.2%; Score 17; DB 4; Length 1070;
Best Local Similarity 100.0%; Pred. No. 81;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1007 CCATGGAGGAGCCGCAG 1023
DB 134 CCATGGAGGAGCCGCAG 150

RESULT 11
US-08-347-792-20
; Sequence 20, Application US/08347792
; Patent No. 5573925
; GENERAL INFORMATION:
; APPLICANT: Halazonetis, Thanos D.
; TITLE OF INVENTION: p53 Proteins with Altered
; TITLE OF INVENTION: Tetramerization Domains
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Howson and Howson
; STREET: Spring House Corporate Cntr., PO Box 457
```

```
; CITY: Spring House
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19477
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentln Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/347,792
; FILING DATE:
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Bak, Mary E.
; REGISTRATION NUMBER: 31,215
; REFERENCE/DOCKET NUMBER: WST580USA
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-540-9206
; TELEFAX: 215-540-5818
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1215 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-347-792-20

Query Match          1.2%; Score 17; DB 1; Length 1215;
Best Local Similarity 100.0%; Pred. No. 81;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1007 CCATGGAGGAGCCGCAG 1023
DB 26 CCATGGAGGAGCCGCAG 42

RESULT 12
US-08-431-357-20
; Sequence 20, Application US/08431357
; Patent No. 5721340
; GENERAL INFORMATION:
; APPLICANT: Halazonetis, Thanos D.
; TITLE OF INVENTION: p53 Proteins with Altered
; TITLE OF INVENTION: Tetramerization Domains
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Howson and Howson
; STREET: Spring House Corporate Cntr., PO Box 457
; CITY: Spring House
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19477
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentln Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/431,357
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/347,792
; FILING DATE: 28-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Bak, Mary E.
; REGISTRATION NUMBER: 31,215
; REFERENCE/DOCKET NUMBER: WST580USA
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-540-9206
; TELEFAX: 215-540-5818
```

INFORMATION FOR SEQ ID NO: 20:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1215 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-08-431-357-20

Query Match 1.2%: Score 17; DB 1; Length 1215;  
Best Local Similarity 100.0%; Pred. No. 81;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1007 CCATGAGAGAGCCGCAG 1023  
|||||  
DB 26 CCATGAGAGAGCCGCAG 42

RESULT 13  
US-08-392-542-26  
Sequence 26, Application US/08392542  
Patent No. 6169073  
GENERAL INFORMATION:

APPLICANT: Halazonetis, Thanos  
APPLICANT: Hartwig, Wolfgang  
TITLE OF INVENTION: Peptides and Peptidomimetics with  
STRUCTURAL SIMILARITY TO HUMAN P53  
TITLE OF INVENTION: Function  
NUMBER OF SEQUENCES: 35  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Banner, Birch, McKie & Beckett  
STREET: 1001 G Street, N.W.  
CITY: Washington, D.C.  
STATE: District of Columbia  
COUNTRY: U.S.  
ZIP: 20001

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/392,542  
FILING DATE:

CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Posorske, Laurence H.  
REGISTRATION NUMBER: 34,698  
REFERENCE/DOCKET NUMBER: 0486.48439  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202 508-9100  
TELEFAX: 202 508-9299

INFORMATION FOR SEQ ID NO: 26:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1215 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA  
ORIGINAL SOURCE:  
ORGANISM: Homo sapiens  
US-08-392-542-26

Query Match 1.2%: Score 17; DB 4; Length 1215;  
Best Local Similarity 100.0%; Pred. No. 81;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1007 CCATGAGAGAGCCGCAG 1023  
|||||  
DB 26 CCATGAGAGAGCCGCAG 42

RESULT 14

US-08-894-327-26  
Sequence 26, Application US/08894327  
Patent No. 6245886  
GENERAL INFORMATION:

APPLICANT: Halazonetis, Thanos  
APPLICANT: Hartwig, Wolfgang  
TITLE OF INVENTION: Peptides and Peptidomimetics with  
STRUCTURAL SIMILARITY TO HUMAN P53 THAT ACTIVATE P53  
TITLE OF INVENTION: Function  
FILE REFERENCE: 2973, 19998  
CURRENT APPLICATION NUMBER: US/08/894,327  
CURRENT FILING DATE: 1997-12-04

EARLIER APPLICATION NUMBER: pctus96/01535  
EARLIER FILING DATE: 1996-02-16  
EARLIER APPLICATION NUMBER: 08/392,542  
EARLIER FILING DATE: 1995-02-16  
NUMBER OF SEQ ID NOS: 35  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 26  
LENGTH: 1215  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-08-894-327-26

Query Match 1.2%: Score 17; DB 4; Length 1215;  
Best Local Similarity 100.0%; Pred. No. 81;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1007 CCATGAGAGAGCCGCAG 1023  
|||||  
DB 26 CCATGAGAGAGCCGCAG 42

RESULT 15  
US-09-685-027-26  
Sequence 26, Application US/09685027  
Patent No. 6420118  
GENERAL INFORMATION:

APPLICANT: Halazonetis, Thanos  
APPLICANT: Hartwig, Wolfgang  
TITLE OF INVENTION: Peptides and Peptidomimetics with  
STRUCTURAL SIMILARITY TO HUMAN P53 THAT ACTIVATE P53  
NUMBER OF SEQUENCES: 35  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Banner, Birch, McKie & Beckett  
STREET: 1001 G Street, N.W.  
CITY: Washington, D.C.  
STATE: District of Columbia  
COUNTRY: U.S.  
ZIP: 20001

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/685,027  
FILING DATE: 10-Oct-2000  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/392,542  
FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:  
NAME: Posorske, Laurence H.  
REGISTRATION NUMBER: 34,698  
REFERENCE/DOCKET NUMBER: 0486.48439  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202 508-9100  
TELEFAX: 202 508-9299

INFORMATION FOR SEQ ID NO: 26:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1215 base pairs

TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA  
ORIGINAL SOURCE:  
ORGANISM: Homo sapiens  
SEQUENCE DESCRIPTION: SEQ ID NO: 26;  
US-09-685-027-26

Query Match 1.2%; Score 17; DB 4; Length 1215;  
Best Local Similarity 100.0%; Pred. No. 81;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1007 CCATGAGAGAGCCGAG 1023  
Db 26 CCATGAGAGAGCCGAG 42

RESULT 16  
PCT-US95-15353-20

Sequence 20, Application PC/TUS9515353  
GENERAL INFORMATION:  
APPLICANT: The Wistar Institute of Anatomy  
APPLICANT: and Biology  
APPLICANT: Halazonetis, Thanos D.  
TITLE OF INVENTION: p53 proteins with altered  
TITLE OF INVENTION: Telomerization Domains  
NUMBER OF SEQUENCES: 46  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Howson and Howson  
STREET: Spring House Corporate Cntr., PO Box 457  
CITY: Spring House  
STATE: Pennsylvania  
COUNTRY: USA  
ZIP: 19477  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/15353  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/347,792  
FILING DATE: 28-NOV-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/431,357  
FILING DATE: 28-APR-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/456,623  
FILING DATE: 01-JUN-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Bak, Mary E.  
REGISTRATION NUMBER: 31,215  
REFERENCE/DOCKET NUMBER: MST58CPCPT  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 215-540-9206  
TELEFAX: 215-540-5818  
INFORMATION FOR SEQ ID NO: 20:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1215 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
PCT-US95-15353-20

Query Match 1.2%; Score 17; DB 5; Length 1215;  
Best Local Similarity 100.0%; Pred. No. 81;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1007 CCATGAGAGAGCCGAG 1023  
Db 26 CCATGAGAGAGCCGAG 42

RESULT 17  
US-08-047-041A-13  
Sequence 13, Application US/08047041A  
Patent No. 5527676

GENERAL INFORMATION:  
APPLICANT: Vogelstein, Bert  
APPLICANT: Baker, Suzanne J.  
APPLICANT: Fearon, Eric R.  
APPLICANT: Nigro, Janice M.  
TITLE OF INVENTION: Detection of loss of the wild-type p53  
NUMBER OF SEQUENCES: 28  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Banner & Allegretti, Ltd.  
STREET: 1001 G Street, N.W.  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20001.4597  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/047,041A  
FILING DATE: 22-MAR-1993  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/928,661  
FILING DATE: 17-AUG-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/446,584  
FILING DATE: 06-DEC-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/330,566  
FILING DATE: 29-MAR-1989  
ATTORNEY/AGENT INFORMATION:  
NAME: Kagan, Sarah A.  
REGISTRATION NUMBER: 32,141  
REFERENCE/DOCKET NUMBER: 01107.42917  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-508-9299  
TELEFAX: 202-508-9100  
INFORMATION FOR SEQ ID NO: 13:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1303 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: Homo sapiens  
POSITION IN GENOME:  
MAP POSITION: 17p13.1  
PUBLICATION INFORMATION:  
AUTHORS: Harris, N.  
JOURNAL: Mol. Cell. Biol  
VOLUME: 6  
ISSUE: 12  
PAGES: 4650-4656  
DATE: 1986

Query Match 1.2%; Score 17; DB 1; Length 1303;  
Best Local Similarity 100.0%; Pred. No. 81;

Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1007 CCATGGAGAGCCGCGAG 1023  
Db 120 CCATGGAGAGCCGCGAG 136

## RESULT 18

US-08-047-041A-12  
Sequence 12, Application US/08047041A  
Patent No. 5527676  
GENERAL INFORMATION:  
APPLICANT: Vogelstein, Bert  
APPLICANT: Baker, Suzanne J.  
APPLICANT: Fearon, Eric R.  
APPLICANT: Nigro, Janice M.  
TITLE OF INVENTION: Detection of Loss of the Wild-Type p53  
NUMBER OF SEQUENCES: 28  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Banner & Allegretti, Ltd.  
STREET: 1001 G Street, N.W.  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20001.4597  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/047,041A  
FILING DATE: 22-MAR-1993  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/928,661  
FILING DATE: 17-AUG-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/446,584  
FILING DATE: 06-DEC-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/330,566  
FILING DATE: 29-MAR-1989  
ATTORNEY/AGENT INFORMATION:  
NAME: Kagan, Sarah A.  
REGISTRATION NUMBER: 32,141  
REFERENCE/DOCKET NUMBER: 01107.42917  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-508-9100  
TELEFAX: 202-508-9299  
INFORMATION FOR SEQ ID NO: 12:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1307 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: Homo sapiens  
POSITION IN GENOME:  
MAP POSITION: 17p13.1  
PUBLICATION INFORMATION:  
AUTHORS: Harris, N.  
TITLE: Molecular basis for heterogeneity of the  
JOURNAL: Mol. Cell. Biol.  
VOLUME: 6  
ISSUE: 12  
PAGES: 4650-4656  
DATE: 1986

US-08-047-041A-12

Query Match 1.2%; Score 17; DB 1; Length 1307;  
Best Local Similarity 100.0%; Pred. No. 81;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1007 CCATGGAGAGCCGCGAG 1023  
Db 124 CCATGGAGAGCCGCGAG 140

## RESULT 19

US-07-912-011-1  
Sequence 1, Application US/07912011  
Patent No. 5382510  
GENERAL INFORMATION:  
APPLICANT: Levine, Arnold J.  
APPLICANT: Shenk, Thomas E.  
APPLICANT: Finlay, Cathy A.  
TITLE OF INVENTION: Probes for Detecting Mutant p53  
NUMBER OF SEQUENCES: 1  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Imclone Systems Incorporated  
STREET: 180 Varick Street  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10014  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/912,011  
FILING DATE: 10-JUL-1992  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Felt, Irving N.  
REGISTRATION NUMBER: 28,601  
REFERENCE/DOCKET NUMBER: LEV-1-C  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-645-1405  
TELEFAX: 212-645-2054  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1317 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
US-07-912-011-1

Query Match 1.2%; Score 17; DB 1; Length 1317;  
Best Local Similarity 100.0%; Pred. No. 81;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1007 CCATGGAGAGCCGCGAG 1023  
Db 134 CCATGGAGAGCCGCGAG 150

## RESULT 20

US-08-347-792-1  
Sequence 1, Application US/08347792  
Patent No. 5573925  
GENERAL INFORMATION:  
APPLICANT: Halazonetis, Thanos D.  
TITLE OF INVENTION: p53 Proteins With Altered  
JOURNAL: Tetramerization Domains  
VOLUME: 37  
ISSUE: 37  
PAGES: 4650-4656  
DATE: 1986  
ADDRESSEE: Howson and Howson

STREET: Spring House Corporate Cntr., PO Box 457  
CITY: Spring House  
STATE: Pennsylvania  
COUNTRY: USA  
ZIP: 19477  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA: US/08/347,792  
APPLICATION NUMBER: US/08/347,792  
FILING DATE:  
CLASSIFICATION: 530  
ATTORNEY/AGENT INFORMATION:  
NAME: Bak, Mary E.  
REGISTRATION NUMBER: 31,215  
REFERENCE/DOCKET NUMBER: WST580USA  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 215-540-9206  
TELEFAX: 215-540-5818  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1317 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 136..1314  
US-08-347-792-1

Query Match 1.2%; Score 17; DB 1; Length 1317;  
Best Local Similarity 100.0%; Pred. No. 81;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1007 CCATGAGAGAGCCGCGAG 1023  
|||||  
Db 134 CCATGAGAGAGCCGCGAG 150

Search completed: December 3, 2002, 10:42:46  
Job time : 94 secs



PI Olson EN, Spencer JA,

Human cDNA encoding  
Human cDNA encoding  
Human secreted pro  
Drosophila melanog  
Drosophila melanog  
Rat sequence diffe  
Human cDNA for nov  
Human cancer relat  
Human wound healin  
Human cDNA encodin  
Human cDNA encodin  
Rat cDNA clone P00  
Human cDNA encodin  
DNA encoding novel  
DNA encoding novel  
DNA encoding novel  
DNA encoding novel  
DNA encoding novel  
Human polynucleotid  
Human polynucleotid  
Human Hnf1r-3 pro  
Human secreted pro  
DNA encoding novel  
Human reproductive  
DNA encoding novel  
Human reproductive  
DNA encoding novel  
Human reproductive  
Human reproductive  
DNA encoding novel  
DNA encoding novel  
Human encoding novel  
Human encoding novel  
Human encoding novel  
PCR primer Fps3-B  
Human silent noncous  
Human dehydrogenase  
Human dehydrogenase  
Novel human diagnosi





PF 20-JUL-2001; 2001WO-IB01903. *ok AZPT dts*  
 XX 28-JUL-2000; 2000US-221607P. *3 too late*  
 PR 02-MAY-2001; 2001US-287724P.  
 XX (COMP-) COMPUGEN INC.  
 PA Shoshan A, Wasserman A, Mintz E, Mintz L, Faigler S;  
 PI WPI; 2002-257383/30.  
 DR  
 XX  
 PF New oligonucleotide libraries comprising oligonucleotides which  
 PT selectively hybridize to mRNAs transcribed from a transcription unit of  
 PT a genome, useful for detecting tissue-, pathology-, and  
 PT developmental-specific genes -  
 PS  
 XX Example 1; SEQ ID 25127; 47bp; English.

The present invention describes oligonucleotide libraries for detecting messenger RNAs that populate a (sub-)transcriptome, where the (sub-)transcriptome comprises messenger RNAs transcribed from multiple transcription units that populate a genome. The library comprises several oligonucleotides, each capable of hybridizing selectively to a set of messenger RNAs transcribed from a given transcription unit of the genome, which encodes one or more messenger RNA splice variants. The oligonucleotide libraries are useful for detecting mRNAs from a biological sample. In expression profiling studies, in qualitatively or quantitatively characterizing the corresponding transcriptome, and in detecting RNA transcripts and splice variants of human or animal transcriptomes. The libraries may also be used as specialised mini libraries to detect transcripts of a sub-transcriptome under a particular biological or pathological state, and so allowing the detection of tissue- and pathology-specific genes such as those genes only expressed in specific tissue under a specific pathological condition; to detect developmental specific genes; and to detect RNA transcripts and splice variants of a transcriptome of a patient suffering from a particular disorder. ABN27253 to ABN59589 represent oligonucleotide sequences from rats, humans and mice, which are used in the exemplification of the present invention.

N.B. The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published\_pat\_sequences.

Sequence 65 BP; 17 A; 13 C; 21 G; 14 T; 0 other;

Query Match 4.5%; Score 65; DB 24; Length 65;  
 Best Local Similarity 100.0%; Pred. No. 1.7e-21;  
 Matches 65; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1114 TATGAGCATGAGCAATCTCTGTGAGCGTGAGCAGCGGAAATGTTGGAAC 1173  
 |||||||  
 DB 1 TATGAGCATGAGCAATCTCTGTGAGCGTGAGCAGCGGAAATGTTGGAAC 60

OY 1174 ATCGA 1178  
 |||||  
 DB 61 ATCGA 65

RESULT 3  
 ID AAA72433 standard; cDNA; 1349 BP.  
 XX  
 AC AAA72433;  
 XX

DT 19-DEC-2000 (first entry)

XX Human nucleic acid-binding protein NuABP-52 cDNA.

XX Human nucleic acid-binding protein; NuABP; agonist; antagonist; EST;  
 KM expressed sequence tag; drug screening; recombinant expression; antibody;  
 KM reproductive disorder; infertility; immunological disorder;  
 KM neurological disorder; cell proliferative disorder; cancer; tumour; ss.  
 XX

OS Homo sapiens.

XX WO200044900-A2.

XX 03-AUG-2000. *too late*

PD 03-AUG-2000. *not applicable under ATPA*

PF 28-JAN-2000; 2000WO-US02237. *< not applicable under ATPA*  
 XX 29-JAN-1999; 99US-0117904.  
 PR 29-JAN-1999; 99US-0117905.

XX (INCY-) INCYTE PHARM INC.

XX Tang YT, Lai P, Hillman JL, Yue H, Azimzai Y, Lu AMD, Baughn MR;  
 PI Tran B, Shin LL, Au-Young JI;

XX WPI; 2000-499332/44.  
 DR P-PSDB; AAB21048.

Novel nucleic acid binding proteins, used to identify agonists and antagonists of them, for the treatment of reproductive, immunological, neurological and cell proliferative disorders including cancer -

Claim 4; Page 178-179; 180pp; English.

Sequences AAA72382-A72436 represent cDNAs encoding novel human nucleic acid-binding proteins (NuABPs; AAB20997-B21051). These cDNAs were produced by extension from an appropriate EST (expressed sequence tag) using primers designed using the EST. The invention also relates to expression constructs, host cells and transgenic organisms comprising a human NuABP nucleic acid, recombinant production of the human NuABPs, and antibodies against the human NuABPs, and also to methods of screening modulators of human NuABP activity or expression. The human NuABPs, and their agonists and antagonists are used to treat diseases associated with overexpression or underexpression of functional NuABPs. Human NuABP proteins and nucleotides, and NuABP agonists and antagonists can be used to diagnose, treat and prevent reproductive, immunological, neurological and cell proliferative disorders. Reproductive disorders that may be treated using compositions of the invention include infertility, endometriosis, disruptions of the menstrual cycle and disruptions of spermatogenesis. Immunological disorders that may be treated include AIDS, allergies, and autoimmune disorders such as multiple sclerosis, rheumatoid arthritis, diabetes and systemic lupus erythematosus. Neurological disorders that may be treated include epilepsy, neurodegenerative conditions such as Alzheimer's disease and Parkinson's disease, prion diseases such as Creutzfeldt-Jakob disease, and mental disorders such as schizophrenia. Cell proliferative disorders that may be treated include a wide variety of cancers, and also arteriosclerosis, atherosclerosis, cirrhosis and psoriasis.

Sequence 1349 BP; 328 A; 372 C; 428 G; 221 T; 0 other;

Query Match 2.5%; Score 36; DB 21; Length 1349;  
 Best Local Similarity 100.0%; Pred. No. 2.6e-07;  
 Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1099 CGGCGAGCCAGCGTATGAGACATGAGCAATTC 1134  
 |||||||  
 DB 1029 CGGCGAGCCAGCGTATGAGACATGAGCAATTC 1064

RESULT 4  
 ID AAH68563 standard; cDNA; 1913 BP.  
 XX  
 AC AAH68563;  
 XX

DT 13-SEP-2001 (first entry)

XX Human protein HP03115 coding sequence.

XX Human; gene therapy; tumour; ss.  
 XX

OS Homo sapiens.  
 XX WO200142302-A1.  
 XX 14-JUN-2001. ~~XXXXXX~~ *OK AIPADuke but*  
 PD  
 XX  
 XX  
 XX  
 XX 06-DEC-2000; 2000WO-JP08631. *Pubh3rd*  
 PF  
 XX  
 XX 06-DEC-1999; 99JP-0346863.  
 PR 06-DEC-1999; 99JP-0346864.  
 PR 08-FEB-2000; 2000JP-0031062.  
 PR 10-FEB-2000; 2000JP-0034090.  
 PR 10-FEB-2000; 2000JP-0034091.  
 PR 14-FEB-2000; 2000JP-0035829.  
 PR 14-FEB-2000; 2000JP-0035829.  
 PR 14-MAR-2000; 2000JP-0071161.  
 PR 30-MAY-2000; 2000JP-0160851.  
 XX  
 XX (NISC-) JAPAN SCT & TECHNOLOGY CORP.  
 PA  
 XX  
 XX Kato S, Eguchi C, Saeki M;  
 PI  
 XX WPI: 2001-381646/40.  
 DR P-PSDB: AAG93278.  
 XX  
 XX  
 XX Human protein originated from tumor cell line, applicable as drug,  
 PT reagent for studying intracellular protein networks and protein source  
 PT for drug screening, also encoded cDNA for gene diagnosis and gene  
 PT therapy -  
 XX  
 XX Claim 3; Pages 236-240; 471pp; Japanese.  
 XX  
 CC The present sequence is a human protein coding sequence. The human  
 CC protein, preferably originated from tumor cell line, is applicable as a  
 CC drug, a reagent for studying intracellular protein networks and a protein  
 CC source for screening proteins for binding low molecular weight drugs. The  
 CC human protein coding sequence is useful for gene diagnosis and gene  
 CC therapy, expression vectors and transformant cells for detection of  
 CC ligands and receptors.  
 CC  
 XX Sequence 1913 BP; 446 A; 537 C; 612 G; 318 T; 0 other;  
 SQ  
 Query Match 2.5%; Score 36; DB 22; Length 1913;  
 Best local Similarity 100.0%; Pred. No. 2.6e-07;  
 Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1099 CGGCCGAGCCAGGCTATGAGAGCATGAGCAATTC 1134  
 ID 1203 CGGCCGAGCCAGGCTATGAGAGCATGAGCAATTC 1238  
 DB  
 RESULT 5  
 AAH78026  
 ID AAH78026 standard; cDNA; 1990 BP.  
 XX  
 AC AAH78026;  
 XX  
 XX 26-NOV-2001 (first entry)  
 DT  
 XX  
 XX Nucleotide sequence of human cardiomyopathy associated protein (CAP).  
 DE  
 XX Human; cardiomyopathy associated protein; CAP; myocardial biopsy;  
 KW left ventricle assist device; LVAD; cardiomyopathy;  
 KW ventricular tachyarrhythmia; ss.  
 XX  
 XX Homo sapiens.  
 OS  
 XX  
 XX Key Location/Qualifiers  
 FH 283..1359  
 FT /\*tag= a  
 FT /product= "cardiomyopathy associated protein (CAP)"  
 PD  
 XX WO200162767-A1.

XX  
 PD 30-AUG-2001.  
 XX  
 XX 26-FEB-2001; 2001WO-US05888. *OK for AIPAD*  
 PF  
 XX  
 XX 24-FEB-2000; 2000US-0184825. *OK for AIPAD*  
 PR  
 XX  
 XX (GENE-) GENE LOGIC INC.  
 PA  
 XX  
 XX Bednarik D, Greene J, White M;  
 PI  
 XX WPI: 2001-570621/64.  
 DR P-PSDB: AAG63832.  
 XX  
 XX Nucleic acid encoding a cardiomyopathy associated protein that is  
 PT differentially expressed in human left ventricle assist device (LVAD)  
 PT myocardial biopsy samples, useful for diagnosing, preventing or  
 PT treating cardiomyopathy -  
 XX  
 XX  
 PS Claim 2; Page 49-51; 55pp; English.  
 XX  
 CC The present sequence encodes a human cardiomyopathy associated  
 CC protein (CAP) that is differentially expressed in human left ventricle  
 CC assist device (LVAD) myocardial biopsy samples. The predicted  
 CC isoelectric point of CAP is approximately 12.2. The CAP polynucleotide  
 CC and polypeptide are useful for diagnosing, preventing or treating  
 CC cardiomyopathy. The treatment diminishes the occurrence of at least  
 CC one of the following symptoms associated with cardiomyopathy: reduced  
 CC ejection fraction, increased left ventricular diastolic dimension,  
 CC decreased ventricular wall thickness, increased atrial size, valvular  
 CC regurgitation, exertional intolerance or ventricular tachyarrhythmia.  
 CC Modulators of CAP are also useful for preventing or treating  
 CC cardiomyopathy.  
 CC  
 XX Sequence 1990 BP; 524 A; 535 C; 613 G; 318 T; 0 other;  
 SQ  
 Query Match 2.5%; Score 36; DB 22; Length 1990;  
 Best local Similarity 100.0%; Pred. No. 2.6e-07;  
 Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1099 CGGCCGAGCCAGGCTATGAGAGCATGAGCAATTC 1134  
 ID 1183 CGGCCGAGCCAGGCTATGAGAGCATGAGCAATTC 1218  
 DB  
 RESULT 6  
 ABN85313  
 ID ABN85313 standard; cDNA; 2040 BP.  
 XX  
 AC ABN85313;  
 XX  
 XX 30-SEP-2002 (first entry)  
 DT  
 XX  
 XX Human cytoskeleton-associated protein, CSAP-4, coding sequence.  
 DE  
 XX Human; cytoskeleton-associated protein; CSAP; CSAP-4;  
 KW cell proliferative disorder; viral infection; neurological disorder;  
 KW transgenic animal; antithrombotic; antiproliferative; antiinflammatory;  
 KW vincristine; anticonvulsant; vasodilator; cerebroprotective; neurotropic;  
 KW neuroprotective; cytoskeletal; gene therapy; gene; ss.  
 XX  
 XX Homo sapiens.  
 OS  
 XX  
 XX Key Location/Qualifiers  
 FH 291..1271  
 FT /\*tag= a  
 FT /product= "CSAP-4"  
 PD  
 XX WO200253719-A2.  
 XX  
 XX 11-JUL-2002.  
 PD  
 XX 04-JAN-2002; 2002WO-US00178. *OK AIPADuke*  
 PF

XX 04-JAN-2001; 2001US-260085P. } *not too late*  
 PR 13-FEB-2001; 2001US-268554P.  
 PR 14-FEB-2001; 2001US-269111P.  
 PR 23-FEB-2001; 2001US-271211P.  
 XX (INCY-) INCYTE GENOMICS INC.  
 PA Lu DAM, Baughn MR, Yao MG, Ding L, Honchell CD, Yue H, Tang YT;  
 PI Warren BA, Duggan BM, Xu Y, Walla NK, Griffin JA, Stewart EA;  
 PI Gandhi AR, Khan FM, Thangavelu K, Ison CH, Azimzai Y, Hafalia AJA;  
 PI Gietzen KJ, Lal PG, Sanjanwala MM, Elliott VS;  
 XX WPI: 2002-583611/62.  
 DR P-PSDB: ABB83475.  
 XX Novel isolated human cytoskeleton-associated protein for diagnosing,  
 PT treating or preventing atherosclerosis, psoriasis, leukemia, epilepsy,  
 PT ischemic cerebrovascular disease, cerebral neoplasms and Alzheimer's  
 PT disease -  
 XX Claim 5; Page 154-155; 167pp; English.  
 CC The present sequence is the coding sequence for a human  
 CC cytoskeleton-associated protein (CSAP). The CSAP and its coding sequence  
 CC are useful in the diagnosis, treatment and prevention of a cell  
 CC proliferative disorder such as actinic keratosis, atherosclerosis,  
 CC psoriasis, primary thrombocythaemia, leukaemia; a viral infection such as  
 CC those caused by adenoviruses (acute respiratory disease, pneumonia),  
 CC arenaviruses (lymphocytic choriomeningitis); and a neurological disorder  
 CC such as epilepsy, ischaemic cerebrovascular disease, stroke, cerebral  
 CC neoplasms, Alzheimer's disease, Pick's disease, Huntington's disease or  
 CC amyotrophic lateral sclerosis. The CSAP coding sequence is also useful  
 CC for creating knock out or knock in humanised animals or transgenic  
 CC animals to model human diseases.  
 XX Sequence 2040 BP; 465 A; 584 C; 650 G; 341 T; 0 other;  
 SO  
 XX Query Match 2.5%; Score 35; DB 24; Length 2040;  
 XX Best Local Similarity 100.0%; Pred. No. 2.6e-07;  
 XX Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 1099 CGCGGAGCCAGCTATGAGAGCATGAGCAATTC 1134  
 ||||||||||||||||||||||||||||||||||||||||  
 DB 1095 CGCGGAGCCAGCTATGAGAGCATGAGCAATTC 1130  
 RESULT 7  
 AAF27653  
 ID AAF27653 standard; DNA; 1500 BP.  
 AC AAF27653;  
 XX 30-MAR-2001 (first entry)  
 DT  
 XX DNA encoding human transcriptional regulator protein #22.  
 DE  
 KW Human; transcriptional regulator protein; TXREG; ds.  
 XX Homo sapiens.  
 OS  
 XX MO200078954-A2.  
 PN  
 XX 28-DEC-2000. *too late*  
 PF 15-JUN-2000; 2000MO-US16766. *Not available under ATPA*  
 PR 18-JUN-1999; 99US-0140109.  
 XX (INCY-) INCYTE GENOMICS INC.  
 PA Lal P, Yue H, Tang YT, Baughn MR, Azimzai Y, Tran B;  
 PI

DR WPI: 2001-041425/05.  
 XX Isolated polypeptide with a human transcriptional regulator protein  
 PT sequence is useful for the diagnosis, prevention and treatment of  
 PT disorders associated with the immune, reproductive and cardiovascular  
 PT systems -  
 XX Claim 5; Page 135-136; 142pp; English.  
 CC The present invention relates to human transcriptional regulator  
 CC protein (TXREG) sequences. The antagonist and an agonist of the proteins  
 CC of the invention are used to treat disorders associated with decreased  
 CC or increased expression or activity of TXREG.  
 XX Sequence 1500 BP; 401 A; 378 C; 439 G; 282 T; 0 other;  
 SO  
 XX Query Match 2.4%; Score 35; DB 22; Length 1500;  
 XX Best Local Similarity 100.0%; Pred. No. 8.1e-07;  
 XX Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 685 CTGCCACCATTTACAAAGCCAGAGAGTGAAGT 719  
 ||||||||||||||||||||||||||||||||||||  
 DB 777 CTGCCACCATTTACAAAGCCAGAGAGTGAAGT 811  
 RESULT 8  
 AAC75760  
 ID AAC75760 standard; CDNA; 650 BP.  
 AC AAC75760;  
 XX 08-FEB-2001 (first entry)  
 DT  
 XX Human ORFX ORF1315 polynucleotide sequence SEQ ID NO:2629.  
 DE  
 XX Human; open reading frame; ORFX; detection; cytostatic; hepatotropic;  
 KW vulnery; antiproliferative; antiparkinsonian; neurotropic; neuroprotective;  
 KW immunostimulant; osteopathic; antiallergic; immunosuppressant; cardiant;  
 KW hypotensive; dermatological; immunosuppressive; antidiabetic;  
 KW antiviral; antibacterial; antifungal; antirheumatic; antithyroid;  
 KW antianemic; gene therapy; cancer; proliferative disorder; hypertension;  
 KW neurodegenerative disorder; osteoarthritis; graft vs host disease;  
 KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;  
 KW cholesterol ester storage; systemic lupus erythematosus; infection;  
 KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;  
 KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;  
 KW bone damage; cartilage damage; antiinflammatory disease; coagulation;  
 KW thrombosis; contraceptive; ss.  
 XX Homo sapiens.  
 OS  
 XX MO200058473-A2.  
 PN  
 XX 05-OCT-2000. *too late*  
 PD  
 XX 31-MAR-2000; 2000MO-US08621. *Not available under ATPA*  
 PF  
 XX 31-MAR-1999; 99US-0127607.  
 PR 02-APR-1999; 99US-0127636.  
 PR 05-APR-1999; 99US-0127728.  
 PR 30-MAR-2000; 2000US-0540763.  
 XX (CURA-) CURAGEN CORP.  
 PA Shinkets RA, Leach M;  
 PI WPI: 2000-602362/57.  
 DR P-PSDB: AAB41551.  
 XX Novel nucleic acids and peptides derived from open reading frame X,  
 PT useful for treating e.g. cancers, proliferative disorders,  
 PT neurodegenerative disorders and cardiovascular disease -

XX  
PS  
XX

Claim 5; Page 1873; 5507pp; English

CC which represent the human ORF open reading frames 1 to 3161. The ORF  
CC sequences have activities such as: cytotoxic; hepatotropic; vulnery;  
CC antiparasitic; antiparkinsonian; neurotic; neuroprotective;  
CC osteopathic; anticoagulant; antiarthritic; immunosuppressive;  
CC immunostimulant; cardiant; thrombolytic; coagulant; vasotropic;  
CC antidiabetic; hypotensive; dermatological; immunosuppressive;  
CC antiinflammatory; antibacterial; antiviral; antifungal; antineumatic;  
CC antihypoid; and antinaemic. The sequences can be used for determining  
CC the presence of or predisposition to, or preventing or treating  
CC pathological conditions associated with an ORF-associated disorder. The  
CC nucleic acids can be used to express ORF proteins in gene therapy  
CC vectors. The proteins and nucleic acids may be used to treat cancers,  
CC proliferative disorders, neurodegenerative disorders, osteoarthritis,  
CC graft vs host disease, cardiovascular disease, diabetes mellitus,  
CC hypertension, hypothyroidism, cholesterol ester storage, systemic lupus  
CC erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,  
CC bacterial or fungal infection, malaria, autoimmune disorders, asthma,  
CC allergies, aplastic anaemia, burns, wounds, bone and cartilage damage,  
CC nocturnal haemoglobinuria, antiinflammatory disease; to enhance  
CC coagulation; to inhibit thrombosis; and as a contraceptive.

sequence 650 BP; 171 A; 150 C; 212 G; 116 T; 1 other,

Query Match	2.0%	Score 29:	DB 21:	Length 650:
Best Local Similarity	100.0%	Pred. No.	0.00072:	
Matches 29:	Conservative 0:	Mismatches 0:	Indels 0:	Gaps 0

Dy	211	GTGGGTTTCAAGCCCGCTGCTAGGGATGC	239
Dd	85	GTGGGTTTCAAGCCCGCTAGGGATGC	113

RESULT 9	
ABA99063	
ID	ABA99063 standard; DNA; 1597 BP.

NC	ADA32003;	
XX		
DT	15-JUL-2002	(first entry)
XX		
DE	Murine muscle ring finger protein 3 (MURF-3) coding sequence	
XX		
XX	Muscle ring finger: MURF-3; mouse; cardiant; microtubule;	
KW	intermediate filament; striated muscle; cardiac hypertrophy;	
KW	heart disease; gene; ds.	
XX		
OS	Mus musculus.	

Key	Location/Qualifiers
EH	299.1330
FT	/*tag= a
FT	/product= "MURF-3"
FT	

PN W0200206318-A2

PD 24-JAN-2002

18-JUL-2001; 2001WO-US22896.

18-JUL-2000; 2000US-219020P.

UNIV TEXAS SYSTEM (TEXA)

Olson EN, Spencer JA;

WPI; 2002-241506/29

Novel muscle fiber

Novel muscle ring finger protein useful for drug screening, and for

PT diagnosing and treating diseases, particularly cardiomyopathies  
XX  
PS Claim 4; Page 131-133, 134pp; English.  
XX

The sequence encodes murine muscle ring finger protein 3 (MURF-3). Invention relates to a purified muscle ring finger (MURF) protein, selected from MURF-1, MURF-2 and MURF-3. The MURF proteins of the invention are involved in microtubule and intermediate filament stabilisation of striated muscle cells and have cardiac activity. The MURF proteins are useful for screening a candidate substance for MURF protein-binding activity, in a cell, cell-free system or *in vivo*, and its effect on interaction of MURF with microtubules, homodimerisation of MURF, MURF-1, MURF-2 or MURF-3 stabilisation of microtubules, interaction of MURF with intermediate filaments, e.g. desmin, vimentin and cytokeratin, and heterodimerisation of MURF. The screened compounds are useful for treating and preventing cardiac hypertrophy and heart diseases. MURF proteins are useful as antigens to immunise animals for the production of antibodies.

sequence 1597 BP; 405 A; 404 C; 438 G; 350 T; 0 other;

Query Match	2.0%;	Score 28;	DB 24;	Length 1597;
Best Local Similarity	100.0%;	Pred. No. 0.0022;		
Matches 28; Conservative	0;	Mismatches	0;	Gaps 0;

QY	505	AACATCATTGACATCTACAAGCAGGAGT	532
Db	596	AACATCATTGACATCTACAAGCAGGAGT	623

RESULT 10
AAS26314
ID AAS26314 standard; cDNA; 587 BP

AC	AAS26314;
XX	
DT	07-NOV-2001 (first entry)
vv	

DE Human cDNA encoding a novel secreted protein, Seq ID 493.  
XX  
KW Human: immunosuppressive; antiarthritic; ss; antirheumatic;  
KW cytostatic; cardiant; vasotropic; cerebroprotective; neurotropic;  
KW neuroprotective; antibacterial; virucide; fungicide; ophthalmological;  
KW vulnerrary; secreted protein; rheumatoid arthritis;  
KW hyperproliferative disorder; cardiovascular disorder; cardiac arrest;  
KW cerebrovascular disorder; cerebral ischemia; angiogenesis;  
KW nervous system disorder; Alzheimer's disease; infection; ocular  
KW corneal infection; wound healing; epithelial cell proliferation;  
KW skin ageing; food additive; preservative; antiproliferative.

OS Homo sapiens

PN WO200155322-A2  
XY

PD 02-AUG-2001  
XY

1/-JAN-2001; 2001WO-US01341

PR	04-JAN-2000; 2000US-0179065
PR	04-FEB-2000; 2000US-0180638

PR	24-FEB-2000; 2000US-0184664
PR	02-MAR-2000; 2000TS-0186350

PR 16-MAR-2000; 2000US-0189874  
PR 17-MAR-2000; 2000US-0190076

PR 18-APR-2000; 2000US-0198123

PR 28-JUN-2000: 2000US-0214886

07-JUL-2000: 2000US-0216647

PR 11-JUL-2000: 2000US-0217487

11-JUL-2000; 2000US-0217496

C  
C  
L  
T  
C  
C  
C  
C  
C  
C  
C

100

100

100

one IFA date  
but Beckman is better

PR 14-JUL-2000; 2000US-0218290.  
PR 26-JUL-2000; 2000US-0220963.  
PR 26-JUL-2000; 2000US-0220964.  
PR 14-AUG-2000; 2000US-0224518.  
PR 14-AUG-2000; 2000US-0224519.  
PR 14-AUG-2000; 2000US-0225213.  
PR 14-AUG-2000; 2000US-0225214.  
PR 14-AUG-2000; 2000US-0225266.  
PR 14-AUG-2000; 2000US-0225267.  
PR 14-AUG-2000; 2000US-0225268.  
PR 14-AUG-2000; 2000US-0225270.  
PR 14-AUG-2000; 2000US-0225447.  
PR 14-AUG-2000; 2000US-0225757.  
PR 14-AUG-2000; 2000US-0225758.  
PR 14-AUG-2000; 2000US-0225759.  
PR 18-AUG-2000; 2000US-0226279.  
PR 22-AUG-2000; 2000US-0226681.  
PR 22-AUG-2000; 2000US-0227182.  
PR 23-AUG-2000; 2000US-0227009.  
PR 30-AUG-2000; 2000US-0228924.  
PR 01-SEP-2000; 2000US-0229287.  
PR 01-SEP-2000; 2000US-0229343.  
PR 01-SEP-2000; 2000US-0229344.  
PR 01-SEP-2000; 2000US-0229345.  
PR 05-SEP-2000; 2000US-0229509.  
PR 05-SEP-2000; 2000US-0229513.  
PR 06-SEP-2000; 2000US-0230437.  
PR 06-SEP-2000; 2000US-0230438.  
PR 08-SEP-2000; 2000US-0231242.  
PR 08-SEP-2000; 2000US-0231243.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231413.  
PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0232080.  
PR 12-SEP-2000; 2000US-0232081.  
PR 14-SEP-2000; 2000US-0232397.  
PR 14-SEP-2000; 2000US-0232398.  
PR 14-SEP-2000; 2000US-0232399.  
PR 14-SEP-2000; 2000US-0232400.  
PR 14-SEP-2000; 2000US-0232401.  
PR 14-SEP-2000; 2000US-0233063.  
PR 14-SEP-2000; 2000US-0233064.  
PR 14-SEP-2000; 2000US-0233065.  
PR 21-SEP-2000; 2000US-0234223.  
PR 21-SEP-2000; 2000US-0234274.  
PR 25-SEP-2000; 2000US-0234597.  
PR 25-SEP-2000; 2000US-0234598.  
PR 26-SEP-2000; 2000US-0235484.  
PR 27-SEP-2000; 2000US-0235834.  
PR 27-SEP-2000; 2000US-0235836.  
PR 29-SEP-2000; 2000US-0236337.  
PR 29-SEP-2000; 2000US-0236337.  
PR 29-SEP-2000; 2000US-0236338.  
PR 29-SEP-2000; 2000US-0236339.  
PR 29-SEP-2000; 2000US-0236370.  
PR 02-OCT-2000; 2000US-0236802.  
PR 02-OCT-2000; 2000US-0237037.  
PR 02-OCT-2000; 2000US-0237038.  
PR 02-OCT-2000; 2000US-0237039.  
PR 02-OCT-2000; 2000US-0237040.  
PR 13-OCT-2000; 2000US-0239935.  
PR 13-OCT-2000; 2000US-0239937.  
PR 20-OCT-2000; 2000US-0240960.  
PR 20-OCT-2000; 2000US-0241221.  
PR 20-OCT-2000; 2000US-0241785.  
PR 20-OCT-2000; 2000US-0241786.  
PR 20-OCT-2000; 2000US-0241787.  
PR 20-OCT-2000; 2000US-0241808.  
PR 20-OCT-2000; 2000US-0241809.  
PR 20-OCT-2000; 2000US-0241826.  
PR 01-NOV-2000; 2000US-0244617.  
PR 08-NOV-2000; 2000US-0246474.  
PR 08-NOV-2000; 2000US-0246475.  
PR 08-NOV-2000; 2000US-0246476.  
PR 08-NOV-2000; 2000US-0246477.  
PR 08-NOV-2000; 2000US-0246478.  
PR 08-NOV-2000; 2000US-0246523.  
PR 08-NOV-2000; 2000US-0246524.  
PR 08-NOV-2000; 2000US-0246525.  
PR 08-NOV-2000; 2000US-0246526.  
PR 08-NOV-2000; 2000US-0246527.  
PR 08-NOV-2000; 2000US-0246528.  
PR 08-NOV-2000; 2000US-0246532.  
PR 08-NOV-2000; 2000US-0246609.  
PR 08-NOV-2000; 2000US-0246610.  
PR 08-NOV-2000; 2000US-0246611.  
PR 08-NOV-2000; 2000US-0246613.  
PR 17-NOV-2000; 2000US-0249207.  
PR 17-NOV-2000; 2000US-0249208.  
PR 17-NOV-2000; 2000US-0249209.  
PR 17-NOV-2000; 2000US-0249210.  
PR 17-NOV-2000; 2000US-0249211.  
PR 17-NOV-2000; 2000US-0249212.  
PR 17-NOV-2000; 2000US-0249213.  
PR 17-NOV-2000; 2000US-0249214.  
PR 17-NOV-2000; 2000US-0249215.  
PR 17-NOV-2000; 2000US-0249216.  
PR 17-NOV-2000; 2000US-0249217.  
PR 17-NOV-2000; 2000US-0249218.  
PR 17-NOV-2000; 2000US-0249244.  
PR 17-NOV-2000; 2000US-0249245.  
PR 17-NOV-2000; 2000US-0249264.  
PR 17-NOV-2000; 2000US-0249265.  
PR 17-NOV-2000; 2000US-0249297.  
PR 17-NOV-2000; 2000US-0249299.  
PR 17-NOV-2000; 2000US-0249300.  
PR 01-DEC-2000; 2000US-0250160.  
PR 01-DEC-2000; 2000US-0250391.  
PR 05-DEC-2000; 2000US-0251030.  
PR 05-DEC-2000; 2000US-0251088.  
PR 05-DEC-2000; 2000US-0256719.  
PR 06-DEC-2000; 2000US-0251479.  
PR 08-DEC-2000; 2000US-0251856.  
PR 08-DEC-2000; 2000US-0251868.  
PR 08-DEC-2000; 2000US-0251869.  
PR 08-DEC-2000; 2000US-0251989.  
PR 08-DEC-2000; 2000US-0251990.  
PR 11-DEC-2000; 2000US-0254097.  
PR 05-JAN-2001; 2001US-0259678.  
PA (HUMA-) HUMAN GENOME SCI INC.  
XX  
XX  
PI Rosen CA, Barash SC, Ruben SM;  
XX  
XX MPI: 2001-488783/53.  
DR P-PSDB; AAT016327.  
XX  
XX  
PT New nucleic acid molecules encoding 461 human secreted proteins for  
PT diagnosing, preventing, treating or ameliorating medical conditions and  
PT used as food additives or preservatives -  
PT  
PT  
XX  
XX  
PS Claim 1; SEQ ID NO 493; 980bp; English.  
XX  
XX  
XX The invention relates to isolated nucleic acid molecules and their  
CC encoded secreted proteins. The nucleic acids and proteins are used to  
CC prevent, treat or ameliorate a medical condition in e.g. humans, mice,  
CC rabbits, goats, horses, cats, dogs, chickens or sheep. They  
CC are also used in diagnosing a pathological condition or susceptibility  
CC to a pathological condition. Antibodies to the proteins can also  
CC be used in alleviating symptoms associated with the disorders and in  
CC diagnostic immunoassays e.g. radioimmunoassays or enzyme linked  
CC immunoassay assays (ELISA). Disorders which are diagnosed or treated  
CC include autoimmune diseases e.g. rheumatoid arthritis,  
CC hyperproliferative disorders e.g. neoplasms of the breast or liver,





Db 316 GTGGCTTCAAGCCGCTGCTAGCGGA 341

## RESULT 13

ABL24238/C  
ID ABL24238 standard; DNA; 2620 BP.

XX ABL24238;

XX 26-MAR-2002 (first entry)

XX Drosophila melanogaster genomic polynucleotide SEQ ID NO 24187.

XX Drosophila; developmental biology; cell signalling; insecticide;

XX pharmaceutical; gene; ds.

XX Drosophila melanogaster.

XX WO200171042-A2.

XX 27-SEP-2001.

XX 23-MAR-2001; 2001WO-US09231.

XX 23-MAR-2000; 2000US-191637P.

XX 11-JUL-2000; 2000US-0614150.

XX (PEKE ) PE CORP NY.

XX Venter JC, Adams M, Li PWD, Myers EW;

XX WPI; 2001-656860/75.

XX New isolated nucleic acid detection reagent for detecting 1000 or more

XX genes from Drosophila and for elucidating cell signalling and cell-cell

XX interactions -

XX Claim 1; SEQ ID NO 24187; 21pp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent

XX capable of detecting 1000 or more genes from Drosophila. The invention is

XX useful in developmental biology and in elucidating cell signalling and

XX cell-cell interactions in higher eukaryotes for the development of

XX insecticides, therapeutics and pharmaceutical drugs. The invention

XX discloses genomic DNA sequences (AB116176-AB130511), expressed DNA

XX sequences (AB101840-AB116175) and the encoded proteins

XX (AB57737-AB72072).

XX The sequence data for this patent did not form part of the printed

XX specification, but was obtained in electronic format directly from WIPO

XX at ftp.wipo.int/pub/published\_pcl\_sequences.

XX Sequence 2620 BP; 625 A; 598 C; 595 G; 802 T; 0 other;

XX Query Match 1.4%; Score 20; DB 23; Length 2620;

XX Best Local Similarity 100.0%; Pred. No. 18;

XX Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 472 GTCTATGCGCTGCGCGGAA 491

XX 985 GTCTATGCGCTGCGCGGAA 966

XX Db 985 GTCTATGCGCTGCGCGGAA 966

XX 26-MAR-2002 (first entry)

XX Drosophila melanogaster expressed polynucleotide SEQ ID NO 38711.

XX Drosophila melanogaster expressed polynucleotide SEQ ID NO 38711.

KW Drosophila; developmental biology; cell signalling; insecticide;

KW pharmaceutical; gene; ss.

XX Drosophila melanogaster.

XX WO200171042-A2.

XX 27-SEP-2001.

XX 23-MAR-2001; 2001WO-US09231.

XX 23-MAR-2000; 2000US-191637P.

XX 11-JUL-2000; 2000US-0614150.

XX (PEKE ) PE CORP NY.

XX Venter JC, Adams M, Li PWD, Myers EW;

XX WPI; 2001-656860/75.

XX P-PDB; ABB70640.

XX New isolated nucleic acid detection reagent for detecting 1000 or more

XX genes from Drosophila and for elucidating cell signalling and cell-cell

XX interactions -

XX Claim 1; SEQ ID NO 38711; 21pp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent

XX capable of detecting 1000 or more genes from Drosophila. The invention is

XX useful in developmental biology and in elucidating cell signalling and

XX cell-cell interactions in higher eukaryotes for the development of

XX insecticides, therapeutics and pharmaceutical drugs. The invention

XX discloses genomic DNA sequences (AB116176-AB130511), expressed DNA

XX sequences (AB101840-AB116175) and the encoded proteins

XX (AB57737-AB72072).

XX The sequence data for this patent did not form part of the printed

XX specification, but was obtained in electronic format directly from WIPO

XX at ftp.wipo.int/pub/published\_pcl\_sequences.

XX Sequence 8574 BP; 1935 A; 2333 C; 2424 G; 1882 T; 0 other;

XX Query Match 1.4%; Score 20; DB 23; Length 8574;

XX Best Local Similarity 100.0%; Pred. No. 17;

XX Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 872 AGGAGCGCAAGCGCGACTG 891

XX 3041 AGGAGCGCAAGCGCGACTG 3060

XX Db 3041 AGGAGCGCAAGCGCGACTG 3060

XX Sequence 8574 BP; 1935 A; 2333 C; 2424 G; 1882 T; 0 other;

XX Query Match 1.4%; Score 20; DB 23; Length 8574;

XX Best Local Similarity 100.0%; Pred. No. 17;

XX Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 872 AGGAGCGCAAGCGCGACTG 891

XX 3041 AGGAGCGCAAGCGCGACTG 3060

XX Db 3041 AGGAGCGCAAGCGCGACTG 3060

XX 26-MAR-2002 (first entry)

XX Drosophila melanogaster expressed polynucleotide SEQ ID NO 38708.

XX Drosophila melanogaster expressed polynucleotide SEQ ID NO 38708.



XX (PEKE ) PE CORP NY.  
 PT Venter JC, Adams M, Li PWD, Myers EW;  
 PI WPI: 2001-6556860/75.  
 XX P-PDB; ABB70639.  
 DR  
 XX New isolated nucleic acid detection reagent for detecting 1000 or more  
 PT genes from Drosophila and for elucidating cell signalling and cell-cell  
 PT interactions -  
 XX  
 XX Claim 1; SEQ ID NO 38708; 21bp + Sequence Listing; English.  
 CC The invention relates to an isolated nucleic acid detection reagent  
 CC capable of detecting 1000 or more genes from Drosophila. The invention is  
 CC useful in developmental biology and in elucidating cell signalling and  
 CC cell-cell interactions in higher eukaryotes for the development of  
 CC insecticides, therapeutics and pharmaceutical drugs. The invention  
 CC discloses genomic DNA sequences (AB16176-AB130511), expressed DNA  
 CC sequences (AB101840-AB16175) and the encoded proteins  
 CC (AB57737-AB572072).  
 CC The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences.  
 CC  
 XX Sequence 13239 BP; 3468 A; 3272 C; 3140 G; 3359 T; 0 other;  
 SQ  
 Query Match 1.4%; Score 20; DB 23; Length 13239;  
 Best Local Similarity 100.0%; Pred. No. 17;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 872 AGGACGCAAGGCGCACTG 891  
 DB 7067 AGGACGCAAGGCGCACTG 7048  
 ID ABR62174/C  
 ID ABR62174 standard; cDNA; 488 BP.  
 XX  
 AC ABR62174;  
 XX  
 DT 18-JUN-2002 (first entry)  
 XX  
 DE Rat sequence differentially expressed in response to a hepatotoxin #81.  
 XX  
 KW Rat; ss; hepatotoxin; expressed sequence tag; EST; drug screening;  
 KW differential expression; centrilobular necrosis; steatosis.  
 XX  
 OS Rattus norvegicus.  
 XX  
 PN WO200210453-A2.  
 PD  
 XX 07-FEB-2002.  
 XX  
 PE 30-JUL-2001; 2001WO-US23872.  
 XX  
 PR 31-JUL-2000; 2000US-222040P.  
 PR 02-NOV-2000; 2000US-244880P.  
 PR 11-MAY-2001; 2001US-290029P.  
 PR 15-MAY-2001; 2001US-290645P.  
 PR 22-MAY-2001; 2001US-292336P.  
 PR 06-JUN-2001; 2001US-295798P.  
 PR 13-JUN-2001; 2001US-297457P.  
 PR 19-JUN-2001; 2001US-298844P.  
 PR 09-JUL-2001; 2001US-303459P.  
 XX  
 PA (GENE-) GENE LOGIC INC.  
 XX  
 PI Mendrick D, Porter MW, Johnson KR, Castle AL, Elashoff MR;  
 XX WPI: 2002-241625/29.  
 DR

XX  
 PT Predicting toxic effects of compounds or the progression of these toxic  
 PT effects by determining the changes in gene expression in tissues or  
 PT cells exposed to the toxin and comparing these to gene expression in  
 PT unexposed tissues or cells -  
 XX  
 XX Claim 1; Seq ID No 81; 239bp; English.  
 PS  
 XX  
 CC The invention relates to methods for predicting toxic effects of  
 CC compounds or the progression of these toxic effects by determining the  
 CC global changes in gene expression in tissues or cells exposed to the  
 CC toxin and comparing these to gene expression in unexposed tissues or  
 CC cells. Also included are methods of predicting at least one toxic  
 CC effect of a compound or progression of a toxic effect, preferably the  
 CC hepatotoxicity of a compound, comprising detecting the level of  
 CC expression in a tissue or cell sample exposed to the compound of two or  
 CC more genes listed in the specification, where differential expression of  
 CC the genes is indicative of at least one toxic effect or progression.  
 CC The method can also be used to identify an agent which modulates the  
 CC toxic response and predict cellular pathways that a compound modulates  
 CC in a cell. The methods utilise a set of at least two probes (on a solid  
 CC support in kit form), where each of the probes comprises a sequence that  
 CC specifically hybridises to a gene listed in the specification, a computer  
 CC system comprising a database containing information identifying the  
 CC expression level in a tissue or cell sample exposed to a hepatotoxin of a  
 CC set of genes comprising at least two genes listed in the specification,  
 CC and a user interface to view the information used to present information,  
 CC identifying the expression level in a tissue or cell of at least one gene  
 CC listed in the specification. The method is useful for elucidating global  
 CC changes in gene expression and for identifying toxicity markers in  
 CC tissues or cell exposed to a known toxin. The genes may be used as  
 CC toxicity markers in drug screening and toxicity assays. The genes and  
 CC gene expression information may be used as diagnostic markers for the  
 CC prediction or identification of the physiological state of tissue or cell  
 CC sample that has been exposed to a compound or agent. Hepatotoxicity  
 CC is characterised by centrilobular necrosis and steatosis. The present  
 CC sequence is an expressed sequence tag (EST) or cDNA derived from a gene  
 CC which is differentially expressed in response to a hepatotoxic agent.  
 CC  
 XX  
 SQ Sequence 488 BP; 93 A; 136 C; 136 G; 123 T; 0 other;  
 Query Match 1.3%; Score 19; DB 24; Length 488;  
 Best Local Similarity 100.0%; Pred. No. 57;  
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1197 TGGGATGAGAGATGAC 1215  
 DB 390 TGGGATGAGAGATGAC 372  
 ID ABR34528  
 ID ABR34528 standard; cDNA; 596 BP.  
 XX  
 AC ABR34528;  
 XX  
 DT 08-MAY-2002 (first entry)  
 XX  
 DE Human cDNA for novel secreted protein, SPQ ID 297.  
 XX  
 KW Human; ss; gene; secreted protein; immune deficiency; viral infection;  
 KW bacterial infection; fungal infection; autoimmune disorder; burn;  
 KW rheumatoid arthritis; multiple sclerosis; autoimmune thyroiditis;  
 KW diabetes; allergy; asthma; neurodegenerative disease; Crohn's disease;  
 KW Alzheimer's disease; Parkinson's disease; liver fibrosis; tumor;  
 KW coagulation disorder; hemophilia; inflammatory disorder; ulcer;  
 KW tissue regeneration; wound healing; hematopoiesis; myeloid deficiency;  
 KW lymphoid cell deficiency.  
 XX  
 OS Homo sapiens.  
 XX  
 PI WO200177290-A2.  
 XX

PD 18-OCT-2001.  
 XX  
 PF 29-MAR-2001; 2001WO-US10295.  
 XX  
 PR 06-APR-2000; 2000US-194941P.  
 XX  
 PA (GEMV ) GENETICS INST INC.  
 XX  
 PI Wong GG, Clark HF, Feschel K, Agostino MJ, Howes SH, Resnick RJ,  
 PI Gulukota K, Graham JR;  
 XX  
 DR WPI: 2002-179323/23.  
 XX  
 PT Six hundred and twenty five polynucleotides derived from a variety of  
 PT human tissue sources which encode secreted proteins, useful for  
 PT treating immune deficiencies and disorders such as autoimmune disorders  
 XX  
 PS Claim 1; Page 152; 339pp; English.  
 XX  
 CC The invention relates to 625 polynucleotides which have been derived from  
 CC a variety of human tissue sources and which encode novel secreted  
 CC proteins, their complements and sequences that hybridise to them.  
 CC Also included are a vector comprising the polynucleotide, a host cell  
 CC transformed with the vector, the proteins encoded by the  
 CC polynucleotides, antibodies that bind to the proteins and identification  
 CC of modulators of the proteins or the expression of the polynucleotide.  
 CC The polynucleotides can be used as probes for the identification  
 CC and isolation of full length cDNA and genomic DNA. The polynucleotides  
 CC and proteins can also be used as nutritional supplements. The protein  
 CC is useful in the treatment of various immune deficiencies and disorders  
 CC such as viral infections, bacterial infections, fungal infections,  
 CC autoimmune disorders (e.g. rheumatoid arthritis, multiple sclerosis,  
 CC conditions (e.g. asthma). They are also useful for treating  
 CC neurodegenerative diseases (e.g. Alzheimer's disease, Parkinson's  
 CC disease), liver fibrosis, coagulation disorders (e.g. haemophilia),  
 CC inflammatory disorders (e.g. Crohn's disease) and tumours. They are also  
 CC useful for tissue regeneration, for wound healing and in the treatment  
 CC of burns, incisions and ulcers. The proteins are also useful for  
 CC regulating haematopoiesis, for treating myeloid or lymphoid cell  
 CC deficiencies. The present sequence is one of the 625 cDNA sequences  
 CC encoding a secreted protein.  
 XX  
 SQ Sequence 596 BP; 116 A; 152 C; 235 G; 93 T; 0 other;  
 XX  
 Query Match 1.3%; Score 19; DB 24; Length 596;  
 Best Local Similarity 100.0%; Pred. No. 57;  
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1247 CGGGCGTGGAGGAGGACG 1265  
 ||||||||||||||||  
 DB 250 CGGGCGTGGAGGAGGACG 268  
 ||||||||||||||||  
 RESULT 18  
 ID ABA82018/c  
 XX ABA82018 standard; cDNA; 629 BP.  
 XX  
 AC ABA82018;  
 XX  
 DT 28-JUN-2002 (first entry)  
 XX  
 DE Human cancer related polynucleotide SEQ ID NO 408.  
 XX  
 KW Human; cytosolic; gene expression; gene mapping; tissue profiling;  
 KW gene therapy; cancer; tumour; gene; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PA WO200214500-A2.  
 XX  
 PN 21-FEB-2002.  
 XX  
 DR

XX  
 PF 16-AUG-2001; 2001WO-US25840.  
 XX  
 PR 16-AUG-2000; 2000US-226326P.  
 XX  
 PA (CHIR ) CHIRON CORP.  
 PA (HXSE-) HXSED INC.  
 XX  
 PI Escobedo J, Garcia PD, Sudduth-Klinger J, Reinhard C, Randazzo F,  
 PI Lamson G, Scott EM, Zhang G, Kassam A, Pot D, Labat I;  
 XX  
 DR WPI: 2002-241905/29.  
 XX  
 PT New nucleic acid for producing a polypeptide, detecting differentially  
 PT expressed genes correlated with a cancerous state of a mammalian cell,  
 PT and inhibiting tumor growth -  
 XX  
 PS Claim 1; SEQ ID NO 408; 883pp + Sequence Listing; English.  
 XX  
 CC The invention relates to an isolated polynucleotide (ABA827253-ABA833262)  
 CC with cytosolic activity. The polynucleotide is used to produce a  
 CC polypeptide, to detect differentially expressed genes correlated with a  
 CC cancerous state of a mammalian cell and to inhibit tumour growth. The  
 CC polynucleotide is used as a probe in mapping and tissue profiling. The  
 CC encoded polypeptide and antibodies to the polypeptide can also be used  
 CC for therapeutic and diagnostic purposes. The polynucleotide is useful for  
 CC gene therapy.  
 CC Note: The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences.  
 XX  
 SQ Sequence 629 BP; 172 A; 195 C; 90 G; 172 T; 0 other;  
 XX  
 Query Match 1.3%; Score 19; DB 24; Length 629;  
 Best Local Similarity 100.0%; Pred. No. 57;  
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1080 GAAGGTGGAGCTGGCAGGA 1098  
 ||||||||||||||||  
 DB 253 GAAGGTGGAGCTGGCAGGA 235  
 ||||||||||||||||  
 RESULT 19  
 ID ABA82018/c  
 XX ABA82018 standard; cDNA; 637 BP.  
 XX  
 AC ABA82018;  
 XX  
 DT 25-JAN-2002 (first entry)  
 XX  
 DE Human wound healing related polynucleotide SEQ ID NO 97.  
 XX  
 KW Human; mouse; vulnery; dermatological; skin disorder; wound healing;  
 KW gene therapy; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PA CA2325226-A1.  
 XX  
 PN 17-MAY-2001.  
 XX  
 PD 16-NOV-2000; 2000CA-2325226.  
 XX  
 PF 17-NOV-1999; 99DE-1055349.  
 XX  
 PR 17-DEC-1999; 99US-0172511.  
 XX  
 PR 20-JUN-2000; 2000DE-1030149.  
 XX  
 PA (SWIT-) SWITCH BIOTECH AG.  
 XX  
 PI Regenbogen J, Wolf E, Goppelt A, Werner S, Halle J;  
 XX  
 DR WPI: 2001-433142/47.  
 XX

PT Use of novel polypeptide or its variant or nucleic acid encoding the  
 PT polypeptide for diagnosing and/or preventing and/or treating skin  
 PT disorders and/or treatment in wound healing or for identifying active  
 PT substances -  
 PS Disclosure; Page 243-244; 265pp; English.  
 XX  
 CC The invention relates to the use of a polypeptide (ABB44544-ABB44601,  
 CC ABB44606-ABB44623) or its variant or encoding nucleic acid  
 CC (ABA81990-ABA81995, ABA82016-ABA82032) with vulnerary and/or  
 CC dermatological activity for the diagnosis, prevention and treatment of  
 CC skin disorders and treatment in wound healing or for the identification  
 CC of pharmacologically active substances. The nucleic acids are useful in  
 CC gene therapy.  
 CC Note: The printed sequence listing for this specification was incomplete,  
 CC terminating part way through SEQ ID NO 106. The remaining data was  
 CC obtained from EPO data for an equivalent patent (EP1114862).  
 XX  
 SQ Sequence 637 BP; 131 A; 184 C; 127 G; 195 T; 0 other;  
 Query Match 1.3%; Score 19; DB 22; Length 637;  
 Best Local Similarity 100.0%; Pred. No. 57;  
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 908 GGGAGCAGGAGGAGCAATT 926  
 Db 62 GGGAGCAGGAGGAGCAATT 44  
 RESULT 20  
 AAS22581  
 ID AAS22581 standard; cDNA; 737 BP.  
 XX  
 AC AAS22581;  
 XX  
 DT 24-OCT-2001 (first entry)  
 XX  
 DE Human cDNA encoding a novel human protein #147.  
 XX  
 KW Human; novel protein; ss; Antinaemic; osteopathic; antiinflammatory;  
 KW immunomodulatory; cytostatic; neuroprotective; vulnerary; nootropic;  
 KW anticonvulsant; antiarthritic; cerebroprotective; antifungal; antiviral;  
 KW antibacterial; antiallergic; dermatological; haemostatic; antiasthmatic;  
 KW thrombolytic; immunogen; antibody; gene therapy; neurological disorder;  
 KW Parkinson's disease; inflammatory disorder; cancer; asthma; osteoporosis;  
 KW tissue regeneration; immune disorder.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200155437-A2.  
 XX  
 PD 02-AUG-2001.  
 XX  
 PF 25-JAN-2001; 2001WO-US02623.  
 XX  
 PR 25-JAN-2000; 2000US-0491404.  
 XX  
 PA (HYSE-) HYSEQ INC.  
 XX  
 PI Tang YT, Liu C, Dirmanac RT;  
 XX  
 DR WPI: 2001-451939/48.  
 DR P-PSDB: AAU14276.  
 XX  
 PT Isolated polypeptides useful for treating anti-inflammatory diseases,  
 PT nervous system disorders, and for regenerating bone and cartilage -  
 PS Claim 1; Page 375-376; 894pp; English.  
 XX  
 CC The invention relates to polynucleotides encoding novel human  
 CC proteins or their active domains. The polypeptides, polynucleotides and  
 CC antibodies raised against the polypeptides are used in a method of  
 CC treatment of a mammal and prevention of disorders caused by the aberrant

CC protein expression or activity. The polypeptides can be used as  
 CC molecular weight markers, food supplements, and in antibody production.  
 CC The polypeptides are used to identify compounds which bind to the  
 CC polypeptides. Polynucleotides of the invention are used as probes and  
 CC primers, for sequencing, for chromosome or gene mapping, in the  
 CC production of recombinant proteins, and in generating anti-sense DNA or  
 CC RNA and in gene therapy. Polypeptides of the invention can be used to  
 CC target drugs to a tumour, in assays to determine biological activity, to  
 CC raise antibodies/elicite an immune response, to determine quantitative  
 CC protein levels, as tissue markers, and to isolate receptors or ligands.  
 CC Polypeptides of the invention may also be useful in treating platelet  
 CC disorders, stem cell disorders, regenerating bone, cartilage, tendon,  
 CC ligament and/or nerve tissue, wound healing, treating burns, promoting  
 CC the proliferation, differentiation and survival of stem cells, as a  
 CC contraceptive, treating osteoporosis and osteoarthritis, anaemia,  
 CC Alzheimer's, Parkinson's and Huntington's diseases, amyotrophic lateral  
 CC sclerosis, stroke, immune deficiencies resulting from bacterial, viral or  
 CC fungal infection or from autoimmunity, cancer, allergy, asthma,  
 CC graft-versus-host disease, eczema, haemophilia, thrombosis,  
 CC anti-inflammatory diseases, nervous system disorders, and infection.  
 CC The present sequence encodes a protein of the invention.  
 XX  
 SQ Sequence 737 BP; 246 A; 111 C; 146 G; 234 T; 0 other;  
 Query Match 1.3%; Score 19; DB 22; Length 737;  
 Best Local Similarity 100.0%; Pred. No. 57;  
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 1115 ATGAGAGCATGAGCAATT 1133  
 Db 699 ATGAGAGCATGAGCAATT 717

Search completed: December 3, 2002, 10:41:33  
 Job time : 309 secs

up C  
D 2A  
E 100  
F 100  
G 100  
H 100  
I 100  
J 100  
K 100  
L 100  
M 100  
N 100  
O 100  
P 100  
Q 100  
R 100  
S 100  
T 100  
U 100  
V 100  
W 100  
X 100  
Y 100  
Z 100

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 3, 2002, 10:41:43 ; Search time 65 Seconds  
(without alignments)  
8478.465 Million cell updates/sec

Title: US-09-908-988B-1

Perfect score: 1431  
Sequence: 1 aagagagtgtacagagagtgtc.....ataagagactcaagtgtccc 1431

Scoring table: OLIGO\_NUC  
Gapop 60.0 , Gapext 60.0

Searched: 341543 seqs, 192557720 residues

Word size : 15

Total number of hits satisfying chosen parameters: 832

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: listing first 45 summaries

Database :

**Published Applications - NA: \***  
1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq: \*  
2: /cgn2\_6/ptodata/1/pubpna/PCR\_NEW\_PUB.seq: \*  
3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq: \*  
4: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq: \*  
5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq: \*  
6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq: \*  
7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq: \*  
8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq: \*  
9: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq: \*  
10: /cgn2\_6/ptodata/1/pubpna/US05\_PUBCOMB.seq: \*  
11: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq: \*  
12: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq: \*  
13: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq: \*  
14: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1431	100.0	1431	10	US-09-908-988B-1
2	28	2.0	1597	10	US-09-908-988B-5
3	27	1.9	587	10	US-09-764-864-493
4	27	1.9	1762	10	US-09-764-864-34
5	19	1.3	488	10	US-09-917-800A-81
6	19	1.3	2513	10	US-09-822-849A-268
7	18	1.3	399	10	US-09-960-352-207
8	18	1.3	444	10	US-09-960-352-13152
9	18	1.3	467	9	US-09-917-800A-577
10	18	1.3	467	9	US-09-736-457-753
11	18	1.3	746	10	US-09-822-849A-580
12	18	1.3	860	10	US-09-833-381-303
13	18	1.3	3033	12	US-10-044-090-508
14	18	1.3	4153	10	US-09-917-800A-1593
15	18	1.3	684973	10	US-09-263-959-1
16	17	1.2	133	10	US-09-776-695-22
17	17	1.2	192	10	US-09-878-574-9448
18	17	1.2	230	10	US-09-864-761-22504
19	17	1.2	230	10	US-09-864-761-22504

20	17	1.2	231	10	US-09-923-876-2132	Sequence 2132, Ap
21	17	1.2	262	10	US-09-878-574-8578	Sequence 8578, Ap
22	17	1.2	264	10	US-09-294-093B-2808	Sequence 2808, Ap
23	17	1.2	264	10	US-09-764-864-499	Sequence 499, Ap
24	17	1.2	274	10	US-09-923-876-5150	Sequence 5150, Ap
25	17	1.2	286	10	US-09-878-574-10927	Sequence 10927, A
26	17	1.2	292	10	US-09-294-093B-1300	Sequence 1300, Ap
27	17	1.2	295	10	US-09-294-093B-1812	Sequence 1812, Ap
28	17	1.2	347	10	US-09-864-761-22849	Sequence 22849, A
29	17	1.2	370	10	US-09-960-352-2716	Sequence 2716, Ap
30	17	1.2	417	10	US-09-867-701-4631	Sequence 4631, Ap
31	17	1.2	446	10	US-09-764-864-482	Sequence 482, Ap
32	17	1.2	459	10	US-09-864-761-31683	Sequence 31683, A
33	17	1.2	465	10	US-09-864-761-6101	Sequence 6101, Ap
34	17	1.2	485	10	US-09-864-761-5742	Sequence 5742, Ap
35	17	1.2	497	10	US-09-864-761-15156	Sequence 15156, A
36	17	1.2	498	10	US-09-783-590-8999	Sequence 8999, Ap
37	17	1.2	535	10	US-09-974-300-3014	Sequence 3014, Ap
38	17	1.2	570	10	US-09-864-761-14784	Sequence 14784, A
39	17	1.2	588	9	US-09-768-020-75	Sequence 75, Appl
40	17	1.2	774	10	US-09-764-864-40	Sequence 40, Appl
41	17	1.2	815	10	US-09-818-954A-18	Sequence 18, Appl
42	17	1.2	990	10	US-09-974-300-523	Sequence 523, Ap
43	17	1.2	1041	10	US-09-778-844-129	Sequence 129, Ap
44	17	1.2	1043	10	US-09-778-844-130	Sequence 130, Ap
45	17	1.2	1048	10	US-09-822-849A-117	Sequence 117, Ap

#### ALIGNMENTS

RESULT 1  
US-09-908-988B-1  
Sequence 1, Application US/09908988B  
Patent No. US20020127690A1  
GENERAL INFORMATION:  
APPLICANT: OLSON, ERIC  
TITLE OF INVENTION: SPENCER, JEFFREY A.  
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR STABILIZING MICROTUBULES  
FILE REFERENCE: MYOG:02805  
CURRENT APPLICATION NUMBER: US/09/908,988B  
CURRENT FILING DATE: 2000-07-18  
PRIOR APPLICATION NUMBER: 60/219,020  
PRIOR FILING DATE: 2000-07-18  
NUMBER OF SEQ ID NOS: 6  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO. 1  
LENGTH: 1431  
TYPE: DNA  
ORGANISM: Mus musculus  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (199)..(1296)  
US-09-908-988B-1  
Query Match 100.0%; Score 1431; DB 10; Length 1431;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1431; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
1 AAGCAGGTGACAGAGGTGCTGGAATATAGACAGGGGTGAGAGGAGCTGTAGGGAAG 60  
|||||  
1 AAGCAGGTGACAGAGGTGCTGGAATATAGACAGGGGTGAGAGGAGCTGTAGGGAAG 60  
|||||  
61 GACAGGACTCTTCCAGAGGAGGAGCAATAGCCGGATCCCAAGATCCAGTACGCTTAAC 120  
|||||  
61 GACAGGACTCTTCCAGAGGAGGAGCAATAGCCGGATCCCAAGATCCAGTACGCTTAAC 120  
|||||  
121 TGACCGAGGAGAGGTGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 180  
|||||  
121 TGACCGAGGAGAGGTGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 180  
|||||  
181 GCTCCAGAGGCGCCGGGAGTGAAGTCAAGGCTGAGGCTGCTAGGGAGTGC 240  
|||||

```
Db 181 GCTCCAGAGCGCGGGAGTGAAGTTCAGCGTGGTTCAAGCGGCTGCTAGGGGATGCG 240
Qy 241 CACACATGAGACACTTGGAGAGCAGCTATTTGCCCATCTGCTGGAGATTTCTCC 300
Db 241 CACACATGAGACACTTGGAGAGCAGCTATTTGCCCATCTGCTGGAGATTTCTCC 300
Qy 301 AAGCCCGTGTATCTTGGCCCTGCCAACACACCTGTGGCGCAGTGTGCCAACAGAGCTC 360
Db 301 AAGCCCGTGTATCTTGGCCCTGCCAACACACCTGTGGCGCAGTGTGCCAACAGAGCTC 360
Qy 361 TTCCAGGCTCTTAATCTCTGTGGCAATCCCGGGGCTCCACAAAGTGTCTTCCAGAGGA 420
Db 361 TTCCAGGCTCTTAATCTCTGTGGCAATCCCGGGGCTCCACAAAGTGTCTTCCAGAGGA 420
Qy 421 CTTTCCAGGCTCTTAATCTCTGTGGCAATCCCGGGGCTCCACAAAGTGTCTTCCAGAGGA 480
Db 421 CTTTCCAGGCTCTTAATCTCTGTGGCAATCCCGGGGCTCCACAAAGTGTCTTCCAGAGGA 480
Qy 481 CTGACGCGGAACTGCTAGTGGAGAACATCATGACATCTACAGCAGAGTCTCCCGG 540
Db 481 CTGACGCGGAACTGCTAGTGGAGAACATCATGACATCTACAGCAGAGTCTCCCGG 540
Qy 541 CCACGACGCGGAACTGCTAGTGGAGAACATCATGACATCTACAGCAGAGTCTCCCGG 600
Db 541 CCACGACGCGGAACTGCTAGTGGAGAACATCATGACATCTACAGCAGAGTCTCCCGG 600
Qy 601 AACATCTACTGCTGAGTGGAGGAGTGGCGGCTCTCTCTGCAAGGTTTCCGCGCC 660
Db 601 AACATCTACTGCTGAGTGGAGGAGTGGCGGCTCTCTCTGCAAGGTTTCCGCGCC 660
Qy 661 CACAGAGACTGTGAGAGTGGCGGCTCTCTCTGCAAGGTTTCCGCGCC 720
Db 661 CACAGAGACTGTGAGAGTGGCGGCTCTCTCTGCAAGGTTTCCGCGCC 720
Qy 721 AGCATGAGTGGAGTGGCGGCTCTCTCTGCAAGGTTTCCGCGCC 780
Db 721 AGCATGAGTGGAGTGGCGGCTCTCTCTGCAAGGTTTCCGCGCC 780
Qy 781 ATGAGAGAGTGGAGTGGCGGCTCTCTCTGCAAGGTTTCCGCGCC 840
Db 781 ATGAGAGAGTGGAGTGGCGGCTCTCTCTGCAAGGTTTCCGCGCC 840
Qy 841 CAGAGTGGAGTGGCGGCTCTCTCTGCAAGGTTTCCGCGCC 900
Db 841 CAGAGTGGAGTGGCGGCTCTCTCTGCAAGGTTTCCGCGCC 900
Qy 901 CTGCGCGGAGAGCAGAGAGTGGCGGCTCTCTCTGCAAGGTTTCCGCGCC 960
Db 901 CTGCGCGGAGAGCAGAGAGTGGCGGCTCTCTCTGCAAGGTTTCCGCGCC 960
Qy 961 GACCACTTGGAGTGGCGGCTCTCTCTGCAAGGTTTCCGCGCC 1020
Db 961 GACCACTTGGAGTGGCGGCTCTCTCTGCAAGGTTTCCGCGCC 1020
Qy 1021 CAGATGAGTGGAGTGGCGGCTCTCTCTGCAAGGTTTCCGCGCC 1080
Db 1021 CAGATGAGTGGAGTGGCGGCTCTCTCTGCAAGGTTTCCGCGCC 1080
Qy 1081 AAGTGGAGTGGAGTGGCGGCTCTCTCTGCAAGGTTTCCGCGCC 1140
Db 1081 AAGTGGAGTGGAGTGGCGGCTCTCTCTGCAAGGTTTCCGCGCC 1140
Qy 1141 AGCGTGGAGTGGAGTGGCGGCTCTCTCTGCAAGGTTTCCGCGCC 1200
Db 1141 AGCGTGGAGTGGAGTGGCGGCTCTCTCTGCAAGGTTTCCGCGCC 1200
Qy 1201 GATGAGAGTGGAGTGGCGGCTCTCTCTGCAAGGTTTCCGCGCC 1260
Db 1201 GATGAGAGTGGAGTGGCGGCTCTCTCTGCAAGGTTTCCGCGCC 1260
Qy 1261 GAGCGCTGGAGTGGCGGCTCTCTCTGCAAGGTTTCCGCGCC 1320
Db 1261 GAGCGCTGGAGTGGCGGCTCTCTCTGCAAGGTTTCCGCGCC 1320
```

```
Db 1261 GAGCGCTGGAGTGGCGGCTCTCTCTGCAAGGTTTCCGCGCC 1320
Qy 1321 ACACCCGAAGCGGAGCCAAAGGATGCTGAGATCTGCGGAGACACCGCGCCACCAA 1380
Db 1321 ACACCCGAAGCGGAGCCAAAGGATGCTGAGATCTGCGGAGACACCGCGCCACCAA 1380
Qy 1381 GCTCGGCTTCCCGCCCGGGAAGGTTTCAATTAAGACTCAAGTGTCC 1431
Db 1381 GCTCGGCTTCCCGCCCGGGAAGGTTTCAATTAAGACTCAAGTGTCC 1431
```

## RESULT 2

```
US-09-908-988b-5
Sequence 5, Application US/09908988B
Patent No. US20020127690A1
GENERAL INFORMATION:
APPLICANT: OLSON, ERIC
APPLICANT: SPENCER, JEFFREY A.
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR STABILIZING MICROTUBULES
FILE REFERENCE: MYOG:0280S
CURRENT APPLICATION NUMBER: US/09/908,988B
CURRENT FILING DATE: 2000-07-18
PRIOR APPLICATION NUMBER: 60/219,020
NUMBER OF SEQ ID NOS: 6
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 5
LENGTH: 1597
TYPE: DNA
ORGANISM: Mus musculus
FEATURE:
NAME/KEY: CDS
LOCATION: (299)..(1327)
US-09-908-988b-5
```

```
Query Match 2.0%: Score 28; DB 10: Length 1597;
Best Local Similarity 100.0%; Pred. No. 0.00016;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 505 AACATCTTGGAGTGGCGGCTCTCTCTGCAAGGTTTCCGCGCC 532
Db 596 AACATCTTGGAGTGGCGGCTCTCTCTGCAAGGTTTCCGCGCC 623
```

## RESULT 3

```
US-09-764-864-493
Sequence 493, Application US/09764864
Patent No. US20020132753A1
GENERAL INFORMATION:
APPLICANT: ROSEN ET AL.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: PT223
CURRENT APPLICATION NUMBER: US/09/764,864
CURRENT FILING DATE: 2001-01-17
Prior application data removed - consult PALM or file wrapper
NUMBER OF SEQ ID NOS: 1792
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 493
LENGTH: 587
TYPE: DNA
ORGANISM: Homo sapiens
US-09-764-864-493
```

```
Query Match 1.9%: Score 27; DB 10: Length 587;
Best Local Similarity 100.0%; Pred. No. 0.00052;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 508 ATCATTTGACATCTACAGCAGAGTCC 534
Db 445 ATCATTTGACATCTACAGCAGAGTCC 471
```

131/00  
15/01

RESULT 4  
US-09-764-864-34  
; Sequence 34, Application US/09764864  
; Patent No. US20020132753A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: PT723  
; CURRENT APPLICATION NUMBER: US/09/764,864  
; CURRENT FILING DATE: 2001-01-17  
; Prior application data removed - consult PAIW or file wrapper  
; NUMBER OF SEQ ID NOS: 1792  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 34  
; LENGTH: 1762  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-764-864-34

Query Match 1.9%; Score 27; DB 10; Length 1762;  
Best Local Similarity 100.0%; Pred. No. 0.00052;  
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 508 ATCATTGACATCTACAGCAGAGTCC 534  
|||||  
DB 467 ATCATTGACATCTACAGCAGAGTCC 493

RESULT 5  
US-09-917-800A-81/C  
; Sequence 81, Application US/09917800A  
; Patent No. US20020119462A1  
; GENERAL INFORMATION:  
; APPLICANT: Mendlick, Donna  
; APPLICANT: Porter, Mark  
; APPLICANT: Johnson, Kory  
; APPLICANT: Castile, Arthur  
; APPLICANT: Elashoff, Michael  
; APPLICANT: Gene Logic, Inc.  
; TITLE OF INVENTION: Molecular Toxicology Modeling  
; FILE REFERENCE: 44921-5038-US  
; CURRENT APPLICATION NUMBER: US/09/917,800A  
; CURRENT FILING DATE: 2001-07-31  
; PRIOR APPLICATION NUMBER: US 60/222,040  
; PRIOR FILING DATE: 2000-07-31  
; PRIOR APPLICATION NUMBER: US 60/222,880  
; PRIOR FILING DATE: 2000-11-02  
; PRIOR APPLICATION NUMBER: US 60/290,029  
; PRIOR FILING DATE: 2001-05-11  
; PRIOR APPLICATION NUMBER: US 60/290,645  
; PRIOR FILING DATE: 2001-05-15  
; PRIOR APPLICATION NUMBER: US 60/292,336  
; PRIOR FILING DATE: 2001-05-22  
; PRIOR APPLICATION NUMBER: US 60/295,798  
; PRIOR FILING DATE: 2001-06-06  
; PRIOR APPLICATION NUMBER: US 60/297,457  
; PRIOR FILING DATE: 2001-06-13  
; PRIOR APPLICATION NUMBER: US 60/298,884  
; PRIOR FILING DATE: 2001-06-19  
; PRIOR APPLICATION NUMBER: US 60/303,459  
; PRIOR FILING DATE: 2001-07-09  
; NUMBER OF SEQ ID NOS: 1740  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 81  
; LENGTH: 488  
; TYPE: DNA  
; ORGANISM: Rattus norvegicus  
; FEATURE:  
; OTHER INFORMATION: Genbank Accession No. US20020119462A1 AA818770  
US-09-917-800A-81

Query Match 1.3%; Score 19; DB 10; Length 488;  
Best Local Similarity 100.0%; Pred. No. 6.2;

Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1197 TGGGATCAAGAGTATGNC 1215  
|||||  
DB 390 TGGGATCAAGAGTATGAC 372

RESULT 6  
US-09-822-849A-268  
; Sequence 268, Application US/09822849A  
; Patent No. US20020045170A1  
; GENERAL INFORMATION:  
; APPLICANT: Wong, Gordon G.  
; APPLICANT: Clark, Hilary  
; APPLICANT: Fechtel, Kim  
; APPLICANT: Agostino, Michael J.  
; APPLICANT: Howes, Steven H.  
; APPLICANT: Resnick, Richard J.  
; APPLICANT: Gulikota, Kamalakkar  
; APPLICANT: Graham, James R.  
; APPLICANT: Genetics Institute, Inc.  
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING NOVEL SECRETED PROTEINS  
; FILE REFERENCE: GIN 6403  
; CURRENT APPLICATION NUMBER: US/09/822,849A  
; CURRENT FILING DATE: 2001-09-04  
; PRIOR APPLICATION NUMBER: 60/195,582  
; PRIOR FILING DATE: 2000-04-06  
; NUMBER OF SEQ ID NOS: 598  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 268  
; LENGTH: 2513  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-822-849A-268

Query Match 1.3%; Score 19; DB 10; Length 2513;  
Best Local Similarity 100.0%; Pred. No. 6.3;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 137 CACAGGCGAGGGGAGAGGC 155  
|||||  
DB 1595 CACAGGCGAGGGGAGAGGC 1613

RESULT 7  
US-09-960-352-207  
; Sequence 207, Application US/09960352  
; Patent No. US20020137139A1  
; GENERAL INFORMATION:  
; APPLICANT: Warren, Wesley C.  
; APPLICANT: Tao, Nengbing  
; APPLICANT: Byatt, John C.  
; APPLICANT: Mathialagan, Nagappan  
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND  
; FILE REFERENCE: 16511.006/37-21(10298)C  
; CURRENT APPLICATION NUMBER: US/09/960,352  
; CURRENT FILING DATE: 2001-09-24  
; NUMBER OF SEQ ID NOS: 15112  
; SEQ ID NO 207  
; LENGTH: 399  
; TYPE: DNA  
; ORGANISM: Bos taurus  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (359)  
; OTHER INFORMATION: unsure at all n locations  
; OTHER INFORMATION: Clone ID: 01-LIB34-072-01-E1-A1  
US-09-960-352-207

Query Match 1.3%; Score 18; DB 10; Length 399;  
Best Local Similarity 100.0%; Pred. No. 20;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	795	CCAGACCATTGAGGACAA	812
Db	275	CCAGACCATTGAGGACAA	292

## RESULT 8

```

US-09-960-352-13152
/ Sequence 13152, Application US/09960352
/ Patent No. US20020137139A1
/ GENERAL INFORMATION:
/ APPLICANT: Warren, Wesley C.
/ APPLICANT: Tao, Mengling
/ APPLICANT: Byatt, John C.
/ APPLICANT: Mathialagan, Nagappan
/ TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
/ TITLE OF INVENTION: MUSCLE AND FAT DEPOSITION
/ FILE REFERENCE: 16511.006/37-21(10298)C
/ CURRENT APPLICATION NUMBER: US/09/960,352
/ CURRENT FILING DATE: 2001-09-24
/ NUMBER OF SEQ ID NOS: 15112
/ SEQ ID NO 13152
/ LENGTH: 444
/ TYPE: DNA
/ ORGANISM: Bos taurus
/ OTHER INFORMATION: Clone ID: 56-LIB34-001-Q1-E1-F8
US-09-960-352-13152

```

Query Match	1.3%	Score 18:	DB 10:	Length 444:
Best Local Similarity	100.0%	Pred. No. 20:		
Matches 18:	Conservative 0:	Mismatches 0:	Indels 0:	Gaps 0

QY	795	CCAGACCATTTGAGGACAA	812
Db	266	CCAGACCATTTGAGGACAA	283

RESULT 9  
US-09-917-800A-577

```

Sequence 577, Application US/09917800A
Patent No US200702119462A1
GENERAL INFORMATION:
Applicant: Mendtick, Donna
Applicant: Porter, Mark
Applicant: Johnson, Kory
Applicant: Castle, Arthur
Applicant: Flashoff, Michael
Applicant: Gene Logic, Inc.
TITLE OF INVENTION: Molecular Toxicology Modeling
FILE REFERENCE: 44921-5038-US
CURRENT APPLICATION NUMBER: US/09/917,800A
PRIOR FILING DATE: 2001-07-31
PRIOR APPLICATION NUMBER: US 60/222,040
PRIOR FILING DATE: 2000-07-31
PRIOR APPLICATION NUMBER: US 60/222,880
PRIOR FILING DATE: 2000-11-02
PRIOR APPLICATION NUMBER: US 60/290,029
PRIOR FILING DATE: 2001-05-11
PRIOR APPLICATION NUMBER: US 60/290,645
PRIOR FILING DATE: 2001-05-15
PRIOR APPLICATION NUMBER: US 60/292,336
PRIOR FILING DATE: 2001-05-22
PRIOR APPLICATION NUMBER: US 60/295,798
PRIOR FILING DATE: 2001-06-06
PRIOR APPLICATION NUMBER: US 60/297,457
PRIOR FILING DATE: 2001-06-13
PRIOR APPLICATION NUMBER: US 60/298,884
PRIOR FILING DATE: 2001-06-19
PRIOR APPLICATION NUMBER: US 60/303,459
PRIOR FILING DATE: 2001-07-09
NUMBER OF SEQ ID NOS: 1740
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 577

```

```

:   LENGTH:445
:   TYPE: DNA
:
:   ORGANISM: Rattus norvegicus
:
:   FEATURE:
:
:   OTHER INFORMATION: Genbank Accession No. US2002011962A1 A1028973
US-09-917-800A-577

```

Query Match	1.3%;	Score 18;	DB 10;	Length 445;
Best Local Similarity	100.0%;	Pred. No. 20;		
Matches	18;	Conservative	0;	Mismatches 0;
			Indels	0;
			Gaps	0

QY	1377	CCAAGCTCGGCTTCCGC	1394
Db	212	CCAAGCTCGGCTTCCGC	229

RESULT 10  
US-09-736

```

Sequence 753, Application US/09736457
Patent No. US20020168637A1
GENERAL INFORMATION:
APPLICANT: Wang, Tongtong
APPLICANT: Bangur, Chaitanya S.
APPLICANT: Iodes, Michael A.
APPLICANT: Fanger, Gary
APPLICANT: Vedvick, Tom
APPLICANT: Carter, Darrick
APPLICANT: Retter, Marc
APPLICANT: Mannion, Jane
APPLICANT: Pan, Liqun
APPLICANT: Wang, Aijun
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
FILE REFERENCE: 210121.478615
CURRENT APPLICATION NUMBER: US/09/7736,457
CURRENT FILING DATE: 2000-12-13
NUMBER OF SEQ. ID NOS: 1864
SOFTWARE: FastSeq for Windows Version 3.0
SEQ. ID NO 753
LENGTH: 467
TYPE: DNA
ORGANISM: Homo sapien
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(467)
OTHER INFORMATION: n = A,T,C or G
US-09-736-457-753

```

Query Match	1.3%;	Score 18;	DB 9;	Length 467;
Best Local Similarity	100.0%;	Pred. No. 20;		
Matches	18;	Conservative	0;	Mismatches
			0;	Indels
				Gaps
				0;

QY	636	CTCTCTCTGCAAGGTTT	6533
Db	442	CTCTCTCTGCAAGGTTT	4253

RESULT 11

US-95-0027941-1/53/C  
Sequence 755, Application US-09/0902941  
Patent No. US20020172952A1  
GENERAL INFORMATION:  
APPLICANT: Henderson, Robert A.  
APPLICANT: Wang, Tongtong  
APPLICANT: Watanabe, Yoshihiro  
APPLICANT: Johnson, Jeffrey C.  
APPLICANT: Retter, Marc W.  
APPLICANT: Mermerakts, Margarita  
APPLICANT: Carter, Darrick  
APPLICANT: Fanger, Gary R.  
APPLICANT: Vedvick, Thomas S.  
APPLICANT: Bangur, Chaltanya S.  
APPLICANT: McNabb, Andria



```

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.478C17
; CURRENT APPLICATION NUMBER: US/09/902,941
; CURRENT FILING DATE: 2001-07-10
; NUMBER OF SEQ ID NOS: 2002
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 753
; LENGTH: 467
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc-feature
; LOCATION: 1, 15, 77, 314, 317, 335, 419
; OTHER INFORMATION: n = A,T,C or G
US-09-902-941-753

Query Match          1.3%; Score 18; DB 9; Length 467;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 636 CTCCTCTGCAAGGTTT 653
    |||
Db 442 CTCCTCTGCAAGGTTT 425

RESULT 12
US-09-822-849A-580
; Sequence 580, Application US/09822849A
; Patent No. US20020045170A1
; GENERAL INFORMATION:
; APPLICANT: Wong, Gordon G.
; APPLICANT: Clark, Hilary
; APPLICANT: Fechtel, Kim
; APPLICANT: Agostino, Michael J.
; APPLICANT: Howes, Steven H.
; APPLICANT: Resnick, Richard J.
; APPLICANT: Gulukota, Kamalakara
; APPLICANT: Graham, James R.
; APPLICANT: Genetics Institute, Inc.
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING NOVEL SECRETED PROTEINS
; FILE REFERENCE: GIN 6403
; CURRENT APPLICATION NUMBER: US/09/822,849A
; CURRENT FILING DATE: 2001-09-04
; PRIOR APPLICATION NUMBER: 60/195,582
; PRIOR FILING DATE: 2000-04-06
; NUMBER OF SEQ ID NOS: 598
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 580
; LENGTH: 746
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-822-849A-580

Query Match          1.3%; Score 18; DB 10; Length 746;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 904 GCCCGGAGCAGAGAG 921
    |||
Db 405 GCCCGGAGCAGAGAG 422

RESULT 13
US-09-833-381-303/c
; Sequence 303, Application US/09833381
; Patent No. US20020132090A1
; GENERAL INFORMATION:
; APPLICANT: Robison, Keith E.
; TITLE OF INVENTION: No. US20020132090A1el Nucleic Acid and Protein Homologs
; FILE REFERENCE: 5800-119
; CURRENT APPLICATION NUMBER: US/09/833,381
; CURRENT FILING DATE: 2001-04-11
```

```

; PRIOR APPLICATION NUMBER: 09/516,448
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 2050
; SOFTWARE: FASTSEQ for Windows Version 3.0
; SEQ ID NO 303
; LENGTH: 860
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc-feature
; LOCATION: (1)...(860)
; OTHER INFORMATION: n = A,T,C or G
US-09-833-381-303

Query Match          1.3%; Score 18; DB 10; Length 860;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 275 GCCCCTCTGCTGGAGA 292
    |||
Db 217 GCCCCTCTGCTGGAGA 200

RESULT 14
US-10-044-090-508
; Sequence 508, Application US/10044090
; Patent No. US20020137081A1
; GENERAL INFORMATION:
; APPLICANT: Olga Bandman
; TITLE OF INVENTION: GENES DIFFERENTIALLY EXPRESSED IN VASCULAR TISSUE ACTIVATION
; FILE REFERENCE: PA-0028 US
; CURRENT APPLICATION NUMBER: US/10/044,090
; CURRENT FILING DATE: 2002-01-09
; NUMBER OF SEQ ID NOS: 850
; SOFTWARE: PERL Program
; SEQ ID NO 508
; LENGTH: 3033
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc-feature
; OTHER INFORMATION: Incyte ID No. US20020137081A1 082155CB1
US-10-044-090-508

Query Match          1.3%; Score 18; DB 12; Length 3033;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 274 TGCCCATCTGCTGGAG 291
    |||
Db 355 TGCCCATCTGCTGGAG 372

RESULT 15
US-09-917-800A-1593
; Sequence 1593, Application US/09917800A
; Patent No. US20020119462A1
; GENERAL INFORMATION:
; APPLICANT: Mendrick, Donna
; APPLICANT: Porter, Mark
; APPLICANT: Johnson, Kory
; APPLICANT: Castile, Arthur
; APPLICANT: Elashoff, Michael
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Molecular Toxicology Modeling
; FILE REFERENCE: 44921-5038-US
; CURRENT APPLICATION NUMBER: US/09/917,800A
; CURRENT FILING DATE: 2001-07-31
; PRIOR APPLICATION NUMBER: US 60/222,040
; PRIOR FILING DATE: 2000-07-31
; PRIOR APPLICATION NUMBER: US 60/222,880
; PRIOR FILING DATE: 2000-11-02
; PRIOR APPLICATION NUMBER: US 60/290,029
```

PRIOR FILING DATE: 2001-05-11  
PRIOR APPLICATION NUMBER: US 60/290,645  
PRIOR FILING DATE: 2001-05-15  
PRIOR APPLICATION NUMBER: US 60/292,336  
PRIOR FILING DATE: 2001-05-22  
PRIOR APPLICATION NUMBER: US 60/295,798  
PRIOR FILING DATE: 2001-06-06  
PRIOR APPLICATION NUMBER: US 60/297,457  
PRIOR FILING DATE: 2001-06-13  
PRIOR APPLICATION NUMBER: US 60/298,884  
PRIOR FILING DATE: 2001-06-19  
PRIOR APPLICATION NUMBER: US 60/303,459  
PRIOR FILING DATE: 2001-07-09  
NUMBER OF SEQ ID NOS: 1740  
SOFTWARE: Patentln Ver. 2.1  
SEQ ID NO 1593  
LENGTH: 4153  
TYPE: DNA  
ORGANISM: Rattus norvegicus  
FEATURE:  
OTHER INFORMATION: Genbank Accession No. US20020119462A1 NM\_012999  
US-09-917-800A-1593

Query Match 1.3%; Score 18; DB 10; Length 4153;  
Best Local Similarity 100.0%; Pred. No. 20;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 608 ACTGCTGAGCTGCAGG 625  
|||||  
DB 2643 ACTGCTGAGCTGCAGG 2660

RESULT 16  
US-09-263-959-1/c  
Sequence 1, Application US/09263959  
Patent No. US20020150891A1  
GENERAL INFORMATION:  
APPLICANT: Hood, Leroy E.  
APPLICANT: Rowen, Lee  
TITLE OF INVENTION: DIAGNOSTIC AND THERAPEUTIC COMPOSITIONS AND METHODS WHICH UTI  
NUMBER OF SEQUENCES: 1279  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Seed and Berry LLP  
STREET: 6300 Columbia Center, 701 Fifth Avenue  
CITY: Seattle  
STATE: Washington  
COUNTRY: US  
ZIP: 98104-7092  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: IBM PC compatible  
SOFTWARE: Patentln Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/263,959  
FILING DATE: 05-MAR-1999  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Mcmasters, David D.  
REGISTRATION NUMBER: 33,963  
REFERENCE/DOCKET NUMBER: 920010.426C2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206) 622-4900  
TELEFAX: (206) 682-6031  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 684973 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-09-263-959-1

Query Match 1.3%; Score 18; DB 10; Length 684973;  
Best Local Similarity 100.0%; Pred. No. 20;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 1034 ACTTCACAGCAGCAAG 1051  
|||||  
DB 526448 ACTTCACAGCAGCAAG 526431

RESULT 17  
US-09-776-695-22  
Sequence 22, Application US/09776695  
Patent No. US20020068283A1  
GENERAL INFORMATION:  
APPLICANT: Boeke, Jef  
Brachmann, Rainer  
TITLE OF INVENTION: NUCLEIC ACIDS ENCODING P53  
MUTATIONS WHICH SUPPRESS P53 CANCER MUTA- TIONS  
NUMBER OF SEQUENCES: 32  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Banner & Wilcoff  
STREET: 1001 G Street, NW  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20001

COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
FILING DATE: 26-Feb-2001  
APPLICATION NUMBER: US/09/776,695  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/795,006  
FILING DATE: 2001-03-28  
ATTORNEY/AGENT INFORMATION:  
NAME: Kagan, Sarah A  
REGISTRATION NUMBER: 32141  
REFERENCE/DOCKET NUMBER: 01107.03170  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-508-9100  
TELEFAX: 202-508-9299  
TELEX: <unknown>

INFORMATION FOR SEQ ID NO: 22:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 133 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
SEQUENCE DESCRIPTION: SEQ ID NO: 22:  
US-09-776-695-22

Query Match 1.2%; Score 17; DB 10; Length 133;  
Best Local Similarity 100.0%; Pred. No. 65;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1007 CCATGAGAGCCGCGAG 1023  
|||||  
DB 42 CCATGAGAGCCGCGAG 58

RESULT 18  
US-09-878-574-9448/c  
Sequence 9448, Application US/09878574  
Patent No. US20020110548A1  
GENERAL INFORMATION:  
APPLICANT: Byrum, Joseph R.  
APPLICANT: La Rosa, Thomas J.  
APPLICANT: Thompson, Michael D.  
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with

```

; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(15401)B
; CURRENT APPLICATION NUMBER: US/09/878,574
; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 09/333,535
; PRIOR FILING DATE: 1999-06-14
; NUMBER OF SEQ ID NOS: 15775
; SEQ ID NO 9448
; LENGTH: 192
; TYPE: DNA
; ORGANISM: Glycine max
; OTHER INFORMATION: Clone ID: 701102502H1
US-09-878-574-9448

Query Match          1.2%; Score 17; DB 10; Length 192;
Best Local Similarity 100.0%; Pred. No. 65;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1248 GGGGCTGGAGAGAGAGC 1264
      |||
Db 162 GGGGCTGGAGAGAGAGC 146

RESULT 19
US-09-864-761-22504
; Sequence 22504, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aecmice-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29

```

```

; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 22504
; LENGTH: 230
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC004477.1
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.2
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.4
; OTHER INFORMATION: EXPRESSED IN HEPA, SIGNAL = 2.2
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2.7
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.2
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.1
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.2
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.9
; OTHER INFORMATION: NT HIT: X84060.1, EVALUATE 1.00e-127
; OTHER INFORMATION: SWISSPROT HIT: Q14494, EVALUATE 1.00e-38
; OTHER INFORMATION: EST_HUMAN HIT: BE897190.1, EVALUATE 1.00e-116
US-09-864-761-22504

```

```

Query Match          1.2%; Score 17; DB 10; Length 230;
Best Local Similarity 100.0%; Pred. No. 65;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 487 CGGAACCTGCTAGTGA 503
      |||
Db 166 CGGAACCTGCTAGTGA 182

```

```

RESULT 20
US-09-923-876-2132
; Sequence 2132, Application US/09923876
; Patent No. US20020013958A1
; GENERAL INFORMATION:
; APPLICANT: Laligudi, Raghunath V.
; APPLICANT: Kamigaki, Laura Y. (Ito)
; APPLICANT: Sherman, Bradley K.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN SEEDLING
; FILE REFERENCE: PL-0012-1 CON
; CURRENT APPLICATION NUMBER: US/09/923,876
; CURRENT FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: 09/298,329
; PRIOR FILING DATE: 1999-04-21
; PRIOR APPLICATION NUMBER: 60/085,331
; PRIOR FILING DATE: 1998-05-05
; NUMBER OF SEQ ID NOS: 6332
; SOFTWARE: PERL Program
; SEQ ID NO 2132
; LENGTH: 231
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc.feature
; OTHER INFORMATION: Incyte ID No. US20020013958A1 700160095H1
; LOCATION: 190, 220
; OTHER INFORMATION: a, t, c, g, or other
US-09-923-876-2132

```

```

Query Match          1.2%; Score 17; DB 10; Length 231;
Best Local Similarity 100.0%; Pred. No. 65;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 499 GTGAGAACATCATGA 515
      |||
Db 123 GTGAGAACATCATGA 139

```

```

Search completed: December 3, 2002, 12:14:06
Job time : 535 secs

```



GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 3, 2002, 06:49:10 : Search time 2791 Seconds

(without alignments)  
14921.575 Million cell updates/sec

Title: us-09-908-988b-1

Sequence: 1 aagagagctagacagagatgtt.....ataagagactcaaggtctcc 1431

Scoring table: IDENTITY-NUC  
Gapop 10.0, Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database :

GenEmbl: 1  
1: gb\_pat: 1  
2: gb\_hlg: 1  
3: gb\_in: 1  
4: gb\_ov: 1  
5: gb\_ov: 1  
6: gb\_pat: 1  
7: gb\_pat: 1  
8: gb\_pat: 1  
9: gb\_pat: 1  
10: gb\_pat: 1  
11: gb\_pat: 1  
12: gb\_pat: 1  
13: gb\_pat: 1  
14: gb\_pat: 1  
15: gb\_pat: 1  
16: gb\_pat: 1  
17: gb\_pat: 1  
18: gb\_pat: 1  
19: gb\_pat: 1  
20: gb\_pat: 1  
21: gb\_pat: 1  
22: gb\_pat: 1  
23: gb\_pat: 1  
24: gb\_pat: 1  
25: gb\_pat: 1  
26: gb\_pat: 1  
27: gb\_pat: 1  
28: gb\_pat: 1  
29: gb\_pat: 1  
30: gb\_pat: 1  
31: gb\_pat: 1  
32: gb\_pat: 1  
33: gb\_pat: 1  
34: gb\_pat: 1  
35: gb\_pat: 1  
36: gb\_pat: 1  
37: gb\_pat: 1  
38: gb\_pat: 1  
39: gb\_pat: 1  
40: gb\_pat: 1  
41: gb\_pat: 1

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1431	100.0	1431	6	AX418848
2	1431	100.0	1448	10	AF294790
3	772.4	54.0	1329	9	HSR291714
4	626.4	43.8	1500	6	AX060632
5	484.6	33.9	1597	6	AX418852
6	483.6	33.8	2097	9	HSR291713
7	482	33.7	1756	9	HSR276484
8	482	33.7	1764	6	AX274927
9	482	33.7	1764	9	AX056642
10	482	33.7	1791	9	AF353673
11	476.8	33.3	1861	10	AY059627
12	475.6	33.2	1211	9	BC015717
13	450.8	31.5	2590	6	AX418850
14	450.6	31.5	1750	9	BC007750
15	450.6	31.5	1925	9	AK091728
16	450.6	31.5	2202	9	HSR291712
17	450.6	31.5	2634	9	AK091310
18	447.4	31.3	1810	9	HSR243488
19	447.4	31.3	2098	9	HSR243489
20	367	25.6	209211	2	AC084883
21	350	24.5	1746	9	AB047601
22	267.4	18.7	531	6	AX313116
23	248.6	17.4	1903	9	AF361946
24	237.8	16.6	199296	2	AC114619
25	189.4	13.2	118847	2	AC013413
26	173.4	12.1	49743	2	AC114541
27	162.6	11.4	164772	2	AC095184
28	154	10.8	123280	2	AC025562
29	107.4	7.5	121721	2	AC090195
30	107.4	7.5	183912	2	AC021240
31	107.2	7.5	168672	2	AC128256
32	107.2	7.5	181514	2	AC095643
33	106.8	7.5	176006	9	AL391650
34	106.4	7.4	20092	10	AF627253
35	106.4	7.4	73989	2	AC101573
36	106.4	7.4	196461	10	AL627514
37	105.8	7.4	170980	2	AF277371
38	104.4	7.3	174303	2	AC099104
39	103.8	7.3	160040	2	AC105674
40	103	7.2	174303	2	AC099104
41	100.2	7.0	160040	2	AC105674
42	97	6.8	1875	9	AK026882
43	95.4	6.7	3127	9	AK096158
44	94.6	6.6	168672	2	AC128256
45	90.6	6.3	2716	9	AB056399

## ALIGNMENTS

RESULT 1  
LOCUS AX418848 1431 bp DNA linear PAT 18-JUN-2002  
DEFINITION Sequence 1 from Patent WO0206318.  
ACCESSION AX418848  
VERSION AX418848.1 GI:21523712  
KEYWORDS  
SOURCE  
ORGANISM  
house mouse.  
Mus musculus  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE  
Olson, E.N. and Spencer, J.A.  
1  
Methods and compositions for stabilizing microtubules and  
intermediate filaments in striated muscle cells

JOURNAL Patent: WO 0206318-A 1 24-JAN-2002;  
BOARD OF REGENTS THE UNIVERSITY OF TEXAS SYSTEM (US)  
FEATURES Location/Qualifiers  
source 1. 1431  
/organism="Mus musculus"  
/db\_xref="Caxon:10090"  
199. .1299  
/note="unnamed protein product"

## CDS

/codon\_start=1  
/protein\_id="CAD35455.1"  
/db\_xref="GI:21523713"  
/translation="MNFVGEKPLDGHNMNLEKQLICPLEMFSKPVYLLPCOH  
NLKRCADVQASPLWOSRSTVSSGRCRPSCRHEVLDRIHGVGLPNIIVE  
NIDITYKQESSRPLAKAEHMLMEHEDEKINICSEVEPCLSKYEGAKDEEV  
APLEPIYKROKSELSDDGIAMVAGNDVROAVITOMEVCOITEDNRKQKOLNORFE  
TLCALVERKSELLOALAREDEKLOVRGLIRQYDLESGSKLVESIOSMEROM  
ALYIOAKELINKYAKMSKVELAGRPPEGYEVSVEHVAEMLRITIDPQGAAG  
DEEDDDNALDEBEGNAGLEERLDVPESSGLH"

BASE COUNT 338 a 384 c 473 g 236 t  
ORIGIN

Query Match 100.0%; Score 1431; DB 6; Length 1431;  
Best Local Similarity 100.0%; Pred. No. 1.1e-267;  
Matches 1431; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAGGAGTGTAGACAGAGTGTCTGGAATAAGACAGGGGTAGAGAGAGCTTTAGGGGAAG 60  
DB 1 AAGGAGTGTAGACAGAGTGTCTGGAATAAGACAGGGGTAGAGAGAGCTTTAGGGGAAG 60  
QY 61 GACAGGACTCTTCCAAAGAGGGGAGCAATAGCCGGGATCCCAAGAAATCCAGTCAACCTTAAC 120  
DB 61 GACAGGACTCTTCCAAAGAGGGGAGCAATAGCCGGGATCCCAAGAAATCCAGTCAACCTTAAC 120  
QY 121 TGACCGAGGAAGGAGTGTACAGAGGAGGAGGAAGGCCAAGACAGAGGCCACAGAGGAGCAG 180  
DB 121 TGACCGAGGAAGGAGTGTACAGAGGAGGAGGAAGGCCAAGACAGAGGCCACAGAGGAGCAG 180  
QY 181 GCTCCAGAGCGCCCGGGATGAATCTCAAGGCTTTCAACCGCTGTAGGGGATGCG 240  
DB 181 GCTCCAGAGCGCCCGGGATGAATCTCAAGGCTTTCAACCGCTGTAGGGGATGCG 240  
QY 241 CACAACATGGAACATTTGGAGAGAGCTCAATTTGCCCATCTGCTGGAGATGTTCTCC 300  
DB 241 CACAACATGGAACATTTGGAGAGAGCTCAATTTGCCCATCTGCTGGAGATGTTCTCC 300  
QY 301 AAGCCGCTGGATCTTCCCTGCCAACACAACCTGTGCCGAAGTGTGCCAACGAGCTC 360  
DB 301 AAGCCGCTGGATCTTCCCTGCCAACACAACCTGTGCCGAAGTGTGCCAACGAGCTC 360  
QY 361 TTCAGGCTCTATTCCTGTGGCAATCCGGGGCTCCACAACCGGTCTTCAGAGAGA 420  
DB 361 TTCAGGCTCTATTCCTGTGGCAATCCGGGGCTCCACAACCGGTCTTCAGAGAGA 420  
QY 421 CGTTTCGATGCCCCATCTTGTAGGACAGAGGTTGCTCGACAGGATGTTCTATGTC 480  
DB 421 CGTTTCGATGCCCCATCTTGTAGGACAGAGGTTGCTCGACAGGATGTTCTATGTC 480  
QY 481 CTGACGAGGGAACCTGCTAGTGTGAGAGACATCTTACATCTCAAGCAGAGTCTCCGG 540  
DB 481 CTGACGAGGGAACCTGCTAGTGTGAGAGACATCTTACATCTCAAGCAGAGTCTCCGG 540  
QY 541 CCACCTGAGGCAAGGCTGAACACACCTCATGTGTGAGAGGACAGAGAGAGAGATC 600  
DB 541 CCACCTGAGGCAAGGCTGAACACACCTCATGTGTGAGAGGACAGAGAGAGAGATC 600  
QY 601 AACATCTACTGCTGAGCTGAGAGTCCCACTGCTCTCTGCAAGGTTTTCGGGCC 660  
DB 601 AACATCTACTGCTGAGCTGAGAGTCCCACTGCTCTCTGCAAGGTTTTCGGGCC 660  
QY 661 CACAAGGACTGTAGGTGGCCCTCTGCCACCACTTACAAAGCCAGAGAGTGAAGCTG 720  
DB 661 CACAAGGACTGTAGGTGGCCCTCTGCCACCACTTACAAAGCCAGAGAGTGAAGCTG 720

QY 721 ACCGATGCGATCGCATGCTGTGGTGGCGGCAATGACCGTGTGACAGGCACTGATCACCAG 780  
DB 721 ACCGATGCGATCGCATGCTGTGGTGGCGGCAATGACCGTGTGACAGGCACTGATCACCAG 780  
QY 781 ATGAGAGAGTGTGCGCAGACCATTTAGAGCAACAGCGCCAGACAGAAAGCACTGTTAAAC 840  
DB 781 ATGAGAGAGTGTGCGCAGACCATTTAGAGCAACAGCGCCAGACAGAAAGCACTGTTAAAC 840  
QY 841 CAGAGTTCGAGAGACCTTGTGGCGGTTTGGAGAGCGCAAGGGGCAACTGCTTCAGCA 900  
DB 841 CAGAGTTCGAGAGACCTTGTGGCGGTTTGGAGAGCGCAAGGGGCAACTGCTTCAGCA 900  
QY 901 CTGGCCCGGGAGCAGAGAGAGAAAGTTGAGCGCGTCCGGGCGCTATCCGCCAGTACGGA 960  
DB 901 CTGGCCCGGGAGCAGAGAGAGAAAGTTGAGCGCGTCCGGGCGCTATCCGCCAGTACGGA 960  
QY 961 GACCACCTTGGAGGCGCTCTCAAGCTGTGAGTCCGCGCATCTCACTCCATGGAAGGCGCG 1020  
DB 961 GACCACCTTGGAGGCGCTCTCAAGCTGTGAGTCCGCGCATCTCACTCCATGGAAGGCGCG 1020  
QY 1021 CAGATGCTCTCTACCTTCAGCAGGCAAGAGCTGATCAACAGAGTGGGGCAATGTCG 1080  
DB 1021 CAGATGCTCTCTACCTTCAGCAGGCAAGAGCTGATCAACAGAGTGGGGCAATGTCG 1080  
QY 1081 AAGCTGAGCTGGGAGAGAGCGCGGAGCCAGGCTATGAGACATGAGCAATCTCTGTG 1140  
DB 1081 AAGCTGAGCTGGGAGAGAGCGCGGAGCCAGGCTATGAGACATGAGCAATCTCTGTG 1140  
QY 1141 AGCGTGGAGCAGTGGCGGAATGTTGGCAACCATGCACTTCCAGCCGGGCGCTGCG 1200  
DB 1141 AGCGTGGAGCAGTGGCGGAATGTTGGCAACCATGCACTTCCAGCCGGGCGCTGCG 1200  
QY 1201 GATGAAGAGATGACGACATGCTTTGATGGGAGAGAGGCAATGCGGGCTGGAGAG 1260  
DB 1201 GATGAAGAGATGACGACATGCTTTGATGGGAGAGAGGCAATGCGGGCTGGAGAG 1260  
QY 1261 GAGCGGCTGGAGCTGCCAAGAGGCTGACCTGACCTGACCTGACCTGACCTGACCTGAC 1320  
DB 1261 GAGCGGCTGGAGCTGCCAAGAGGCTGACCTGACCTGACCTGACCTGACCTGACCTGAC 1320  
QY 1321 ACACCCGAAGCGGAGGCAAGGATGCTGAGGATCTGCGACAGACACCGCGCCACCA 1380  
DB 1321 ACACCCGAAGCGGAGGCAAGGATGCTGAGGATCTGCGACAGACACCGCGCCACCA 1380  
QY 1381 GCTCGGCTTCCCGCCCGGGAAGTCTCAATTAAGAGCTCAAGTGTCC 1431  
DB 1381 GCTCGGCTTCCCGCCCGGGAAGTCTCAATTAAGAGCTCAAGTGTCC 1431

RESULT 2  
AF294790 1448 bp mRNA linear ROD 08-SEP-2000  
LOCUS AF294790  
DEFINITION Mus musculus RING-finger protein MURF mRNA, complete cds.  
ACCESSION AF294790  
VERSION AF294790.1 GI:9945009  
KEYWORDS  
SOURCE  
ORGANISM  
Mus musculus.  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 1448)  
REFERENCE  
Spencer, J.A., Eliazzer, S., Ilaria, R.L. Jr., Richardson, J.A. and  
Olson, E.N.  
Regulation of microtubule dynamics and myogenic differentiation by  
MURF, a striated muscle RING-finger protein  
J. Cell Biol. 150 (4), 771-784 (2000)  
JOURNAL  
MEDLINE  
20411220  
PUBMED  
10953002  
REFERENCE  
2 (bases 1 to 1448)  
Spencer, J.A. and Olson, E.N.  
Direct Submission  
TITLE  
Submitted (09-AUG-2000) Molecular Biology, UTSW, 6000 Harry Hines  
Blvd., Dallas, TX 75390-9148, USA

FEATURES  
Source

CDS

Location/Qualifiers  
1..1448  
/organism="Mus musculus"  
/db\_xref="taxon:10090"  
/tissue-type="cardiac muscle"  
199..1299  
/codon\_start=1  
/product="RING-finger protein MURF"  
/protein\_id="AA03076.1"  
/db\_xref="GI:9945010"  
/translation="MNFVGFKPLIGDANMDLEKOLICPLEMFSKPVYLLPCOH  
NLCKKANDVFOASNPILMOSRSTVSSGRCRREVLDRGVGLRNILVE  
NIDITYOESSRPLHAKAHOHMEHEDEKINICLSEVPTSCIKVEGAKDCEV  
APLPTIYKROKSELSDDGAMLVAGNDROAVITOMPEVOTIEDNRKOLINORE  
TLCAVBERKELLDLAREDEKLRVGLRGTGDHLEGGSKVESIQSNEPROM  
ALYLOAKKELINKVAMSKVELAGRPDEGSEMEQFSVEHVAEMLRITIDPQGAAG  
DEEDDDMALDEBNAGLEBERLDVPEBSGLH"

BASE COUNT 355 a 384 c 473 g 236 t  
ORIGIN

Query Match 100.0%; Score 1431; DB 10; Length 1448;  
Best Local Similarity 100.0%; Pred. No. 1,1e-267;  
Matches 1431; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Oy 1 AAGGAGTGTAGACAGAGTGTGTGAATATGACAGGGGTGAGAGAGCTTTAGGGAAG 60
    |||||||
Db 1 AAGGAGTGTAGACAGAGTGTGTGAATATGACAGGGGTGAGAGAGCTTTAGGGAAG 60
    |||||||
Oy 61 GACAGGACTCTTCCAAAGAGGAGCAATAGCCGGATCCCAAGAATCCAGTCAAGCTAAC 120
    |||||||
Db 61 GACAGGACTCTTCCAAAGAGGAGCAATAGCCGGATCCCAAGAATCCAGTCAAGCTAAC 120
    |||||||
Oy 121 TGACCGAGAGAGGTGTCAACAGCGAGGAGAGAGCCCAACAGAGGCGCAAGCGAGCAG 180
    |||||||
Db 121 TGACCGAGAGAGGTGTCAACAGCGAGGAGAGAGCCCAACAGAGGCGCAAGCGAGCAG 180
    |||||||
Oy 181 GGTCCAGAGCGCGCGGGATGAACCTTCAAGCGGTTCAGCGGTCTAGGAGGATCG 240
    |||||||
Db 181 GGTCCAGAGCGCGCGGGATGAACCTTCAAGCGGTTCAGCGGTCTAGGAGGATCG 240
    |||||||
Oy 241 CACACATGATGACAACTTGGAGAGAGAGCTATTTGCCCATCTGCTGAGAGATTTCTCC 300
    |||||||
Db 241 CACACATGATGACAACTTGGAGAGAGAGCTATTTGCCCATCTGCTGAGAGATTTCTCC 300
    |||||||
Oy 301 AAGCCGCTGTATCTTGGCCCTGCAACACACACCTGTGCGCAAGTGTGCCAAGAGATC 360
    |||||||
Db 301 AAGCCGCTGTATCTTGGCCCTGCAACACACACCTGTGCGCAAGTGTGCCAAGAGATC 360
    |||||||
Oy 361 TTCCAGGCTCTATCTTGTGCAATCCGGGGCTCCACACAGATGTCTTCAAGAGAGA 420
    |||||||
Db 361 TTCCAGGCTCTATCTTGTGCAATCCGGGGCTCCACACAGATGTCTTCAAGAGAGA 420
    |||||||
Oy 421 CGTTTCCGATGCCCATCTTGTAGGACAGAGTGTCTCTGCAAGGATGTCTATGGC 480
    |||||||
Db 421 CGTTTCCGATGCCCATCTTGTAGGACAGAGTGTCTCTGCAAGGATGTCTATGGC 480
    |||||||
Oy 481 CTGACAGGAGCACTGTAGTGAAGCAATCATTTGATTCATAGCAGAGATCCGCCG 540
    |||||||
Db 481 CTGACAGGAGCACTGTAGTGAAGCAATCATTTGATTCATAGCAGAGATCCGCCG 540
    |||||||
Oy 541 CCACATGACAGCAAGGCTGTGAACAGACCTCATGTGTAGAGAGCAAGAGAGATC 600
    |||||||
Db 541 CCACATGACAGCAAGGCTGTGAACAGACCTCATGTGTAGAGAGCAAGAGAGATC 600
    |||||||
Oy 601 AACATCTACTGCTGAGCTCGAGAGTCCCACTGCTCTCTGCAAGGTTTGGCGCC 660
    |||||||
Db 601 AACATCTACTGCTGAGCTCGAGAGTCCCACTGCTCTCTGCAAGGTTTGGCGCC 660
    |||||||
Oy 661 CACAAGGAGCTGTGAGTGGCGCTCTGCCCAACATTTACAAGCGCACAAGAGTACG 720
    |||||||
Db 661 CACAAGGAGCTGTGAGTGGCGCTCTGCCCAACATTTACAAGCGCACAAGAGTACG 720
    |||||||
Oy 721 AGGATGATGATCGGATGCTGTGGCGGCAATGACCGTGTGACAGAGTATCACCGAG 780
    |||||||

```

```

Db 721 AGGATGATGATCGGATGCTGTGGCGGCAATGACCGGTGTCAGAGAGTATCACCGAG 780
    |||||||
Oy 781 ATGAGAGAGTGTGTGCGACACATTATAGACAAACACCGCAGACAGAACTGTTAAAC 840
    |||||||
Db 781 ATGAGAGAGTGTGTGCGACACATTATAGACAAACACCGCAGACAGAACTGTTAAAC 840
    |||||||
Oy 841 CAGAGGTCGAGACCCCTGTGCGGTTTGGAGAGCGCAAGGAGCAATGTCTCAAGCA 900
    |||||||
Db 841 CAGAGGTCGAGACCCCTGTGCGGTTTGGAGAGCGCAAGGAGCAATGTCTCAAGCA 900
    |||||||
Oy 901 CTGGCCCGGAGCAGAGAGAGAGAGTTCAGCGCTGCGGGGCTCATCCGCAAGTACGA 960
    |||||||
Db 901 CTGGCCCGGAGCAGAGAGAGAGAGTTCAGCGCTGCGGGGCTCATCCGCAAGTACGA 960
    |||||||
Oy 961 GACCACTTGTGAGAGGCTCTCTCAAAAGTGTGAGTCCGCATCAGTCCATGAGAGCGCG 1020
    |||||||
Db 961 GACCACTTGTGAGAGGCTCTCTCAAAAGTGTGAGTCCGCATCAGTCCATGAGAGCGCG 1020
    |||||||
Oy 1021 CAGATGCTCTCTACCTCTCAGAGAGCAAGAGAGTGTATCAAGAGTGTGCGGCAATGTG 1080
    |||||||
Db 1021 CAGATGCTCTCTACCTCTCAGAGAGCAAGAGAGTGTATCAAGAGTGTGCGGCAATGTG 1080
    |||||||
Oy 1081 AAGTGTGAGCTGTGCGAGAGCGCGGAGCCAGGCTATGAGAGATGAGCAATTTCTGTG 1140
    |||||||
Db 1081 AAGTGTGAGCTGTGCGAGAGCGCGGAGCCAGGCTATGAGAGATGAGCAATTTCTGTG 1140
    |||||||
Oy 1141 AGCGTGTGAGCAAGTGTGCGCAATGTGTGCAACATCAGTCCAGCGGGCGCGGTGG 1200
    |||||||
Db 1141 AGCGTGTGAGCAAGTGTGCGCAATGTGTGCAACATCAGTCCAGCGGGCGCGGTGG 1200
    |||||||
Oy 1201 GATGAAGAGATGATGACATGCTTGTGATGAGGAGAGAGGCAATCGGGCTGTGAGAG 1260
    |||||||
Db 1201 GATGAAGAGATGATGACATGCTTGTGATGAGGAGAGAGGCAATCGGGCTGTGAGAG 1260
    |||||||
Oy 1261 GAGCGCTGTGAGAGTGTGCGAGAGGCTTGTGAGGCTGTGAGCCGACTGTATCCAGAGCG 1320
    |||||||
Db 1261 GAGCGCTGTGAGAGTGTGCGAGAGGCTTGTGAGGCTGTGAGCCGACTGTATCCAGAGCG 1320
    |||||||
Oy 1321 ACACCGAAGGAGGAGGAGCAAGGAGTGTGAGTGTGCGAGAGACCAAGCGCCACCA 1380
    |||||||
Db 1321 ACACCGAAGGAGGAGGAGCAAGGAGTGTGAGTGTGCGAGAGACCAAGCGCCACCA 1380
    |||||||
Oy 1381 GCTGCGCTTCCCGCCCGCGGAGAGTTCATATTAAGAGTCAAGTGTCC 1431
    |||||||
Db 1381 GCTGCGCTTCCCGCCCGCGGAGAGTTCATATTAAGAGTCAAGTGTCC 1431
    |||||||

```

RESULT 3  
HSA291714 1329 bp mRNA linear PRI 06-JUL-2001  
LOCUS HSA291714  
DEFINITION Homo sapiens mRNA for RNF30 gene for ring finger protein 30.  
ACCESSION AJ291714.1 GI:13160387  
VERSION AJ291714.1 GI:13160387  
KEYWORDS MURF-3 gene; muscle specific ring finger protein.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 1329)  
Centner,T., Yano,J., Kimura,E., McElhinny,A.S., Pellin,K.,  
Wilt,C.C., Bang,M.L., Trombitas,K., Granzler,H., Gregorio,C.C.,  
Sotiriachi,H. and Labat,S.  
Identification of the titin kinase domain  
regulators of the titin kinase domain  
J. Mol. Biol. 306 (4), 717-726 (2001)  
JOURNAL MEDLINE  
PUBMED 11243782  
REFERENCE 2 (bases 1 to 1329)  
Centner,T. Structure and Biocomputing,  
Direct Submission  
TITLE JOURNAL  
Submitted (12-FEB-2001) Centner T., Heidelberg 69112, Germany  
EMBL-Heidelberg, Weyherhofstrasse 1,

REMARK Revised by author 22-FEB-2001  
 FEATURES Location/Qualifiers  
 source 1..1329  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 misc\_feature 1..3  
 /note="putative alternative translation initiation codon"  
 /evidence=not\_experimental  
 49..1203  
 /gene="RNF30"  
 49..1203  
 /gene="RNF30"  
 /note="alternative"  
 /codon\_start=1  
 /product="ring finger protein 30"  
 /protein\_id="CAC32842.1"  
 /db\_xref="GI:13160389"  
 /translation="MDNLEKOLICPLEMFSKPVLLPCQNLCKRCANDYFOASNP  
 LMOGRSTTVSSGGRFCPSCRHEVLDLRHGVYQJRNILVENIIDYKQESRPLHS  
 KAOHLMCEHEEEKINITYCLSEVPYSLCKVFCAHKEDEVAIPTIYKROKODLT  
 LLPLRECGTNTYCSLDLPSSDPPIIASONTKIIDPSLGIAMLVANRYQAVI  
 TQMEVCOITIDNSRROKOLLTOREPSICAVIEERKELLOMLAREKELORGLI  
 ROYGDHLEASSKIVESAIOSEMERPOLALYLOQAKELINKVGMASKVELGRPEPYES  
 MEQTVRYEHVAEMLRITDFQPGASGESEEVAPDEBSAGPEERPDGP"  
 join(49..315,642..1203)  
 /gene="RNF30"  
 /note="short isoform"  
 /codon\_start=1  
 /product="ring finger protein 30"  
 /protein\_id="CAC32841.1"  
 /db\_xref="GI:13160388"  
 /translation="MDNLEKOLICPLEMFSKPVLLPCQNLCKRCANDYFOASNP  
 LMOGRSTTVSSGGRFCPSCRHEVLDLRHGVYQJRNILVENIIDYKQESRPLHS  
 KAOHLMCEHEEEKINITYCLSEVPYSLCKVFCAHKEDEVAIPTIYKROKODLT  
 GIAMLVANRYQAVITQMEVCOITIDNSRROKOLLTOREPSICAVIEERKELLO  
 LAREKELORGLIROYGDHLEASSKIVESAIOSEMERPOLALYLOQAKELINKYGA  
 MSKVELGRPEPYESMEQTVRYEHVAEMLRITDFQPGASGESEEVAPDEBSAGP  
 EERPDGP"  
 BASE COUNT 310 a 375 c 414 g 230 t  
 ORIGIN  
 Query Match 54.0%: Score 772.4: DB 9: Length 1329;  
 Best Local Similarity 78.2%: Pred. No. 6.7e-140;  
 Matches 1062: Conservative 0: Mismatches 141: Indels 155: Gaps 5;

QY 619 TCGAGAGTGCCACCTGCTCTCTGCAAGGTTTTCGCGCCACAGAGACTGTGAGTG 678  
 Db 421 TGTGAGGTGCCACCTGCTCTCTGCAAGGTTTTCGCGCCACAGAGACTGTGAGTG 480  
 QY 679 GCCCTTCGCCACCATTTACAAAGCCGCAAG----- 711  
 Db 481 GCCCTTCGCCACCATTTACAAAGCCGCAAGAAAGAGATCTACTGTGCCAGG 540  
 QY 712 ----- 711  
 Db 541 CTGAGTGCAGTGCAGCAACAACACTTACTGACGCTTGATCTCCGAGCTCAAGTAT 600  
 QY 712 ----- AGTAGCTGAGCATGGCATC 732  
 Db 601 CCTCCATCTTACCTCGAGAACACTAAGATTATAGATAGTACCTCAGCATGGCATC 660  
 QY 733 GCGATGCTGTGCGGGGCAATGACCGGTGAGGAGGAGTATACCAAGATGAGAGAGTG 792  
 Db 661 GCGATGCTGTGCGGGGCAATGACCGGTGAGGAGGAGTATACCAAGATGAGAGAGTG 720  
 QY 793 TGCACACCATTTAGAGACACCGCAGACAGAACTGTTAAACCGAGAGTTGAG 852  
 Db 721 TGCACACCATTTAGAGACACCGCAGAGCAAGAGCAAGTGTTAACCAAGAGTTGAG 780  
 QY 853 ACCCTGCGCGGTTTGGAGAGGCGCAAGGCGCAAGTGTTCAGACACTGGCCGGGAG 912  
 Db 781 AGCTGTGCGGAGTGTGAGAGAGGCGCAAGGCGCAAGTGTTCAGACACTGGCCGGGAG 840  
 QY 913 CAGGAGAGAAAGTTGACAGCGGCGTGCAGTCCGCAATGAGAGACACTTGGAG 972  
 Db 841 CAGGAGAGAAAGTTGACAGCGGCGTGCAGTCCGCAATGAGAGACACTTGGAG 900  
 QY 973 GCGTCCCAAGAGCTGTGAGAGTCCGCTTCAAGTCAATGAGAGAGCCGCAAGTGTCTC 1032  
 Db 901 GCGTCCCAAGAGCTGTGAGAGTCCGCTTCAAGTCAATGAGAGAGCCGCAAGTGTCTC 960  
 QY 1033 TACCTCCAGAGCGCAAGAGAGTATCAACAGTGGGAGCAATGTGAGAGTGGAGTG 1092  
 Db 961 TATCTCCAGAGCGCAAGAGAGTATCAACAGTGGGAGCAATGTGAGAGTGGAGTG 1020  
 QY 1093 GCAGAGAGCGCGGAGCGAGCTATGAGAGATGAGAGCAATTTCTGTGAGGTGAGAGC 1152  
 Db 1021 GCAGAGAGCGCGGAGCGAGCTATGAGAGATGAGAGCAATTTCTGTGAGGTGAGAGC 1080  
 QY 1153 GTGGCCCAAGAGTGTGCAACCATGACCTTCAGCGCGCGCGCTGGGAGTGAAGAGAT 1212  
 Db 1081 GTGGCCCAAGAGTGTGCAACCATGACCTTCAGCGCGCGCGCTGGGAGTGAAGAGAT 1140  
 QY 1213 GACGACATGCTTGTGATGAGAGAGAGGAGCAATGCGGGGCTGAGAGAGAGAGTGGAC 1272  
 Db 1141 G-----TGCCCCAGAGCGGAGAGAGAGAGAGAGCGCGGGGCGGAGAGAGAGAGAGT 1194  
 QY 1273 GTGCCAGAAAGCTCAGGCTGACCTGACCGGACTCTGATCCAGAGCGCACCCGAAAGCG 1332  
 Db 1195 GGGCC-----TTAAGGCTCGCGCCGACCGACCTCTGAGAGAGCGCGCTGAGAGTGC 1248  
 QY 1333 GGAGCCAAAGAGTGTGAGAGTGTGCGCAGAGAGACACCGGCGCCAGAGTGTGCTTCC 1392  
 Db 1249 G-----GAGATCTGCGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGTGG---CGC 1291  
 QY 1393 GCCCGCGGAGAGTGTCTCAATTAAGAGACTCAAGTGTCC 1430  
 Db 1292 CGGCCCCGGAGAGTGTCAATTAAGAGACTCAAGTGTCC 1329

RESULT 4  
 AX060632 1500 bp DNA linear PAT 22-JAN-2001  
 LOCUS AX060632  
 DEFINITION Sequence 54 from Patent W00078954.  
 ACCESSION AX060632  
 VERSION AX060632.1 GI:12406053  
 KEYWORDS  
 SOURCE human.



ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE 1 (bases 1 to 1500)  
AUTHORS Lal, P., Yue, H., Tang, Y. T., Baughn, M. R., Azimzal, Y. and Tran, B.  
TITLE Human transcriptional regulator proteins  
JOURNAL Patent: WO 0078954-A 54-28-DEC-2000;  
Incyte Genomics, Inc. (US)

FEATURES  
source 1.1500  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"

BASE COUNT 401 a 378 c 439 g 282 t

ORIGIN

Query Match 43.8%; Score 626.4; DB 6; Length 1500;  
Best Local Similarity 87.5%; Pred. No. 1.4e-111;  
Matches 721; Conservative 0; Mismatches 96; Indels 7; Gaps 3;

Db 118 AACGACCGAGGAGGCTGCACAGCGAG-GGGAGAGGCCAAGACAGCGCCACAGCGAG 176  
214 ATCTAAGCGAGGAGGCTGTACAGCGAGTGAAGGCCAGAGCGCGCCAGGCCAG 273  
177 GGAGGCTCAGAGCGCGGGGATGAAGTTCACGTTGAGGTTTCAAGCGCGTCTAGGGGA 236  
274 GCACGACGCGCGCGGGGATGAAGTTCACAGTGGTTTCAAGCGCGTCTAGGGGA 328

237 TCGCGACAAATGAGCACTGTGAGAGAGAGAGTTCATTTGCCCATCTGCGTGAAGTGT 296  
329 TGCACACACAGTGAAGCACTGTGAGAGAGAGTTCATTTGCCCATCTGCGTGAAGTGT 388  
297 CTCGCAAGCGCGTGTATCTTGGCCCTGCCAACAACCTGCGCGGCAAGTGTGCCAAGA 356  
389 CTCGCAAGCGCGTGTATCTTGGCCCTGCCAACAACCTGCGCGGCAAGTGTGCCAAGA 448

357 CGTCTTCAGAGCGCTTAATCTCTGTGGCAATCCGGGGCTCCCAACAAGGTGTTCAGG 416  
449 CGTCTTCAGAGCGCTTAATCTCTGTGGCAATCCGGGGCTCCCAACAAGGTGTTCAGG 508

417 AGAGAGCTTTCATGCGCCATCTGTGAGGACAGAGTGTCTCTGAGACAGGCAATGGTGT 476  
509 AGGCGCTTTCATGCGCCATCTGTGAGGACAGAGTGTCTCTGAGACAGGCAATGGTGT 568

477 TGGCTTCAGAGCGGAGCTGTGAGAGAGAGTTCATTTGACATTCATCAAGAGAGTCTTC 536  
569 CGGCTTCAGAGCGGAGCTGTGAGAGAGAGTTCATTTGACATTCATCAAGAGAGTCTTC 628

537 CCGGCAAGCGGAGCTGTGAGAGAGAGTTCATTTGACATTCATCAAGAGAGTCTTC 596  
629 CAGGCGGCTGACCTCCAAAGGCTGAGAGAGAGTTCATTTGACATTCATCAAGAGAGT 688

597 GATCAACATCTACTGCTGCTGAGAGAGAGTTCATTTGACATTCATCAAGAGAGTCTTC 656  
689 GATCAACATCTACTGCTGCTGAGAGAGAGTTCATTTGACATTCATCAAGAGAGTCTTC 748

657 CGCCCAAGAGAGCTGTGAGAGAGAGTTCATTTGACATTCATCAAGAGAGTCTTC 716  
749 TGGCCCAAGAGAGCTGTGAGAGAGAGTTCATTTGACATTCATCAAGAGAGTCTTC 808

717 GCTGAGCGAGTGTGCTGCTGAGAGAGAGTTCATTTGACATTCATCAAGAGAGTCTTC 776  
809 GCTGAGCGAGTGTGCTGCTGAGAGAGAGTTCATTTGACATTCATCAAGAGAGTCTTC 868

777 CCGAGGAGAGAGTGTGCTGCTGAGAGAGAGTTCATTTGACATTCATCAAGAGAGTCTTC 836  
869 ACGAGTGTGAGAGAGTGTGCTGCTGAGAGAGAGTTCATTTGACATTCATCAAGAGAGT 928

837 AAACCAAGAGTGTGAGAGAGTGTGCTGCTGAGAGAGAGTTCATTTGACATTCATCAAG 896  
929 AAACCAAGAGTGTGAGAGAGTGTGCTGCTGAGAGAGAGTTCATTTGACATTCATCAAG 988

897 AGCACTGGCGCGGAGAGTGTGCTGCTGAGAGAGAGTTCATTTGACATTCATCAAGAGT 939

Db 989 GCGCGTGGCGCGGAGAACAGCAGGAGCTTCAACGCGATCCG 1032

RESULT 5  
AX418852 1597 bp DNA linear PAT 18-JUN-2002  
LOCUS Sequence 5 from Patent WO0206318.  
DEFINITION  
ACCESSION AX418852  
VERSION AX418852.1 GI:21523716  
KEYWORDS  
SOURCE house mouse.  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1  
AUTHORS Olson, E.N. and Spencer, J.A.  
TITLE Methods and compositions for stabilizing microtubules and intermediate filaments in striated muscle cells  
JOURNAL Patent: WO 0206318-A 5 24-JAN-2002;  
BOARD OF REGENTS THE UNIVERSITY OF TEXAS SYSTEM (US)

FEATURES  
source 1.1597  
/organism="Mus musculus"  
/db\_xref="taxon:10090"  
299.1330  
/note="unnamed protein product"  
/codon\_start=1  
/protein\_id="CAD35457.1"  
/db\_xref="GI:21523717"  
/translation="MDYKSLIPDGNAMENLEKOLICPTCEMTKPPVITLPCOHNLK  
RKANDIRQANPYWTNRGVSMSGRRRCPSHEVIMDHVIGLDRNLVENVIT  
DIYOECSRLPQKSHPMCKEHEDEKIIYCLTEVPYSLCYVGAHQAEVLA  
SIFOGKTELINICIMLVANDRVOITISLQDSCKRKNESHQVKEDELSQKFDLYA  
TIDKESLRLQITQEOEKELGFIALLIYREOEKRSKIVETALIOSLDEFGATFL  
SSAKLIRSVBASGCGQKGTQEGPENNDYITLDEHIAELRAIDFGTKGCDVTC  
LTFEORSS"

BASE COUNT 405 a 404 c 438 g 350 t

ORIGIN

Query Match 33.9%; Score 484.6; DB 6; Length 1597;  
Best Local Similarity 69.2%; Pred. No. 4.6e-84;  
Matches 677; Conservative 0; Mismatches 299; Indels 3; Gaps 1;

Qy 226 CTGCTAGGGGAGTCCGCAACATGACACTTGGAGAGAGCTCATTTGCCCATCTGC 285  
317 CTGATTCCTGATGAGAAAGCTATGAGAGAACTGAGAGAGTCTGCCCATCTGC 376

Db 286 CTGCAATGTCTTCCAAAGCGCGGTGATCTTGCGCCCAACAGCACTGTGCCGAG 345  
377 CTGCAATGTCTTCCAAAGCGCGGTGATCTTGCGCCCAACAGCACTGTGCCGAG 436

Qy 346 TGTGCCAAGAGCTTCTCCAGGCGCTTATCTGTGGCAATCCGGGGCTCCACAG 405  
437 TGTGCCAAGAGCAATCTTCCAGGCGGTGATCTTGCGCCCAACAGCACTGTGCCAG 496

Qy 406 GTGCTTTCAGAGAGAGTTCGCAATGCTGTAGGACAGAGGTGTCTGGAAGAG 465  
497 TCCATGTCTGAGAGTGTCTTCCGCTTGTGCGCGCAATGAGATGATGAGAGAG 556

Db 466 CATGTGTCTATGAGCGCTGCGAGCGGAGCTGTGAGAGAGATCATTTGACATTCAG 525  
557 CAGGAGGTGTAGCGCTGCGAGAGAGAGTGTGAGAGAGATCATTTGACATTCAG 616

Qy 526 CAGAGTCTTCCGGGCACTGCACAGCGGCTGAACAGCACTCATGTGTGAGAGAG 585  
617 CAGAGAGTCTCCAG---TCGGCCCTTCAGAAAGCGAGCCAGCGATGTGCAAGGAGAC 673

Db 586 GAGGAGAGAGATCATCATCTACTGCTGAGCGGAGGAGCGGAGCGGAGCGGAGCGG 645  
674 GAGGAGAGAGATCATCATCTACTGCTGAGCGGAGGAGCGGAGCGGAGCGGAGCGG 733

Qy 646 AAGGTTTCCGCGCCCAAGAGAGTGTGAGTGTGCGCCCTGTGCCACCATTTCAAGAGC 705



Qy	1006	TCATATGAGAGAGCGCAGATNGCTCTCTACTCTCAGACGAGCAAGAGCGATCAACAAG	1065
Db	1231	TTCTTTTGGACAGCCTTGCGGGAGCCACCTTCTCTTGACTGCGCAAGCACTCATCAAAAGC	1290
Qy	1066	GTCGGGGCAATGTCGAAGTGTGAGCTGTGGCAGCAGCGCCGAGCCAGGCTATGAGACATG	1125
Db	1291	ATTGTGGAACCTTCCAAGGGCTCGCAGCTGTGGGAAGACACAGCAGCGCTTTGGAACATG	1350
Qy	1126	GAGCAATTCCTCTGTGAGCGTGTGAGCAGCTGTGGCCGAAATGTTGCGAACCATGCACTTCAG	1185
Db	1351	GACCTTCTTACTTGTGCAATTTAGAGACATATGACAGACGCCCTGTGAGAGCATGTGACTTTGGG	1410
Qy	1186	CCGGGGCCCCCTGGGGATGTGAAGAGATGACACATGGCTTTGGATGGGGAGGGCCAT	1245
Db	1411	ACAGATAGGAAGAAGGAGAAATTCATTATGAAGAAGAAATATGAGAAAGAGAGTCCACA	1470
Qy	1246	GCGGGCGCTGGAGAGAGCAGCGCTGTGACGTGCCAGAAAG	1283
Db	1471	GAGGGAAGGAAGAAGACACACAGTAAGAGAGCTGGATG	1508
RESULT 7			
LOCUS	HS276484		
DEFINITION	HS276484	1756 bp	mRNA
DEFINITION	Human mRNA for muscle specific RING finger 2 (MURF2) protein (MURF2 gene).		linear
ACCESSION	AJ276484		
VERSION	AJ276484.1		
KEYWORDS	MURF2 protein; RING finger protein.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
AUTHORS	1 (bases 1 to 1756)		
TITLE	Characterisation of MURF2, a new muscle-specific RING finger protein of the RBC family that associates with microtubules		
JOURNAL	Unpublished		
REFERENCE	2 (bases 1 to 1756)		
AUTHORS	Stanchi, F.		
TITLE	Direct Submission		
JOURNAL	Submitted (20-MAR-2000) Stanchi F., CRIBI Biotechnology Centre, Universita di Padova, Via G. Colombo 3, Padova, 35121, ITALY		
FEATURES	location/Qualifiers		
SOURCE	1..1756		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/chromosome="1"		
	/map="P35.1-35.2"		
	/clone="STRAIT03036"		
	/tissue_type="skeletal muscle"		
	/clone_lib="pPRK8"		
	/dev_stage="adult"		
	1..1756		
	/gene="MURF2"		
	1..282		
	/gene="MURF2"		
	/number=1		
	124..1185		
	/gene="MURF2"		
	/codon_start=1		
	/product="muscle specific RING finger 2 (MURF2)"		
	/protein_id="CAC81706.1"		
	/db_xref="GI:18073356"		
	/translation="MDYKSSLIQGNPNEINERKOLPICLEMTKRVYILPCQHNLC		
	RKCNADLIFQANPYTSSRGSSVSMGSEFRPTGRHEHYIMRHGVGIQRMLVETITII		
	DIVQOECSRPANPLQSGHPMKCHEDEKINICLICEVPTSCMKVFGIHKACEVAPIQ		
	SVFGQETLENNCSIMLVAGNDRYQTITIOLEDSRYTKENSHOVKELSQKPTLYA		
	ILDKKSELIQRIOEODEKRTSFLEALIOQOEOLODKSTKVERAIOISLDPPGATFL		
	LTAQLIKSIYASKGCOLGKTGTGCFEMDFITLDLHIDALALNADFGTDEDEEFTI		
	EEEOEEESTEGKEEGHO"		
	283..455		
exon			

Query Match	Best Local Similarity	Matches 705;	Conservative	Score 482;	DB 9:	Length 1756;
Matches 705;	Conservative	0;	Mismatches 350;	Indels 3;	Gaps	
QY 226	CTGTAGGAGGAGGCCACACATGAGACACTTGAGAGAGCAGCTATTGGCCCATCTGC	142	CTGTATCCAGAGATGGGAGATGCCATGAGAGACTTGAGAGAGCAGCTGATCTGCTATCTTC	285		
QY 286	CTGAGAGATGTTTCCAGAGCCCGTGGATCTTGGCCCTGCCAACACACCTGTGCGGCAAG	202	CTGGAGATGTTTACCAAGGACAGTGGATCTTGGCCGTGCCACACACACTGTGCGGGAAG	261		
QY 346	TGTGCCAAGCAGCTGCTCCAGGCGCTTAATCCTCTGTGAGCAATCCCGGGGCTCCACAGC	262	TGTGCCAATGACATCTTCCAGGCTGCAATTCCTACTGTGAGCAGCAGCGGGGAGCTGAGT	321		
QY 406	GTTCTTTCAGAGAGCGTTTCCGATGCCATCTTGTAGGACAGAGTTTCTTGACAGG	322	TCCATGTCTGGAGGCGGTTTCCGTGCGTCCGACCTGCCACAGAGGTGATCATGTGATCTGT	381		
QY 466	CATGTTGTTATGCGCTGAGAGGAGCACTGTAGTGGAGAACATCTTATCATCTTACAG	382	CAGGAGTGTAGCGCTGAGAGAGACGTGTGTGTGGAGAACATCTTCGACATCTTACAA	441		
QY 526	CAGAGATCTCTCCCGGCGACTGTCAGGCCAGGCTGAAACAGACACTCATGTGTGAGAGCAC	442	CAGGAGTGTCTCAGTGGCGCGCTGACAGAGGAGGAGT---CACCCCATGTGCAAGGAGCAC	498		
QY 586	GAGAGCAGAAAGATCAACATCTATGCTCTGAGAGTGTGGAGAGTGGCCACACGCTCTCTGC	499	GAGAGTGTGAGAAATCAACATCTATCTCTACGTGTGAGAGTGGCCACGCTCTCATGTGC	558		
QY 646	AAGGTTTGGGCGCCCAAGAGAGCTGTGAGTGGGCGCCCTGTGCCACCATTTTCAAAAGC	559	AAGGTTTGGGAGTTCACAAAGGCTGTGAGTGGGCGCCCATTTGAGAGTGTCTTCCAGGGA	618		
QY 706	CAGAGAGTGTAGCTGTGAGCGATGCGATTCGATCTGTGTGGCGGGCATGACCTGTGACG	619	CAAAAGACTGTGAACTGTGATTACTATCTTCATCTCTGTTGCGGGGAGATGACCTGTGACG	678		
QY 766	GCAGTGTACCCAGAGTGGAGAGGTGTGTCCACACCATTTGAGAGAACAGACCCGACAG	679	ACCATCATCTCAGCTGGAGAGATTCCTCCGTCTGAGTGTACCAAGAGAGAACATCTCACAGGTA	738		

QY	836	AAGCAACTGTTAAACACAGAGCTTTCAGACCCCTGTCGGCGTTTGGAGGAGGCGCAAGGGC	885
Db	739	AAGGAAGAGCTAGCCACAAGATTTCACAGCTGTGTATGCATCTCTGATGAGAGAAAGT	798
QY	886	GAATGCTTTCAAGCACTGGCCCGGGAGCAGAGAGAAAGTTGACAGCCCTGGCGGCTTC	945
Db	799	GAGTGTCTGCACGCGGATCACCGAGGAGCAGGAGAAAGAAAGCTTAGCTTCAATCAGGCGCTTC	858
QY	946	ATCCGCACTACGAGAGACCACTTTGGAGGGCTTCCTAAAGCTGTGAGATGCCCATCCAG	1005
Db	859	ATCCAGGAGTACAGAGAGCAGCTGGACAAAGTCCACAAAGCTGTGGAAACTCCATCCAG	918
QY	1006	TTCATGAGGAGACCAGATGGCTCTCTCACTCCAGCAGGCAAGAGAGCTGATCAACAAG	1065
Db	919	TCCCTGGACGAGCCCTGGGGAGCCACCTTCTTGACTGCCAAGCACTCATCAAAAGC	978
QY	1066	GTCGGGGCAATGTCAAGCTGGAGCTGGCAGAGACGGCCGGAGCCAGGCTATAGAGCATG	1125
Db	979	ATTGGGAAGCTTCCAAAGGCTGCCAGCTGGGGAGACAGACAGGAGCTTTAGAACTTG	1038
QY	1126	GAGCAATTCCTGTAGAGCGTGGAGCAGCTGGCGCAATAGTTTGCAGAACCATGACTTCCAG	1185
Db	1039	GACTCTTTACTTTTGATTTAGAGACACATACGACAGCGCCCTTAGAGCCATTACTTTGGG	1098
QY	1186	CCGGGCCCCGCTGGGGATGAAGAGGATGACGACATGAGCTTTGGATGGGAGAGAGGCAT	1245
Db	1099	ACAAGTGAAGAGAGGAAGAAATTCATTGAAAGAAAGATCAGGAGAGAGAAAGATCCACA	1158
QY	1246	CGGGGCTGGAGAGGAGACGGCTGCAGCCTGCCAGAGG	1283
Db	1159	GAGGGAAGGAAGAACGACACACAGTAAAGAGCTGGATG	1196

LOCUS	AX274927	1764 bp	DNA	linear	PAT 29-OCT-2001
DEFINITION	Sequence 192 from Patent WO0172777.				
ACCESSION	AX274927				
VERSION	AX274927.1	GI:16547559			
KEYWORDS					
SOURCE	human.				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.				
AUTHORS	1 Hillman, J.T., Baughn, M.R., Yue, H., Lal, P., Lu, D.A., Paterson, C., Azimzai, Y., Bandman, O., Tang, Y.T., Mathur, P., Shah, P., Au-Young, J.				
TITLE	Transcription factors				
JOURNAL	Patent: WO 0172777-A 192 04-OCT-2001;				
FEATURES	1. Incyte genomics, Inc. (US) location/Qualifiers				
source	1. .1764 /organism="Homo sapiens" /db_xref="taxon:9606" /note="Incyte ID No: 3575519CB1"				
BASE COUNT	473 a 399 c 516 g 376 t				
ORIGIN					
Query Match	33.7%;	Score 482;	DB 6;	Length 1764;	
Best local Similarity	66.68;	Pred. No. 1.5e-81;			
Matches 705; Conservative	0;	Mismatches 350;	Indels 3;	Gaps 1;	

[illegible]

D	261	TCGCGCAATGACATCTTCCAGCGGTGCAAAATCCCTACTGGACACGCCGGGCAGCTCAGTG	320
Q	406	GTGCTCTCAGGAGAGCGTTCCGATGCCATCTTGTAGGACAGAGTTGTCTGACAGG	465
D	321	TCGATCTCTGGAGGCGGTTCCGCTCCGCCACCTGCGCGCACGAGGATCATGGATCGT	380
Q	466	CATGCGTCTATGCGCTGGACGGACCTGCTGTGTGAGACACATCAATGTGATCTCAAG	525
D	381	CACGGAGTGTACGGCTGCGAGAGAACCTGCTGTGTGAGAACATCTGCAATCTCAAA	440
Q	526	CAGAGTCCCTCCCGCCACTGCACGCCAAGCGTGAACAGCACCTCATGTTGTAGAGAC	585
D	441	CAGGAGTGTCTCCAGTCGCGCGCTGCAGAAAGGCAAT - - - CACCCCATGTGCAAGAGCAC	497
Q	586	GAGGACGGAATATCAACATCTACTGCTGAGCTGCGAGAGTGCACCTGCTCTCTGC	645
D	498	GAAATGAGAAATCAACATCTACGTCTCAGCTGTGAGAGTGCCTCCACTGCTCAATGTG	557
Q	646	AAGTCTTCGCGCGCCACAGAGACTGTGAGTGGCCCTGTGCGCCACATTTACAACGC	705
D	558	AAGTGTTTGGGATTCACAGAGCGCTGCGAGGTGCGCCACTTGACAGTGTCTTCCAGGGA	617
Q	706	CAGAAAGTGTGAGCTGAGCGATGTCATCGGATGCTGTGTGCGCGCAATGACGCTGCGAG	765
D	618	CAAAAGACTGAATCAATTAATCTGTATCTCATGTGTGTGGCGGGAAATGACGCTGTGCGAG	677
Q	766	GCAGTATTCACCAATGAGAGAGGTGTGCCAGACCATTTGAGAGACACGCCGACAGAC	825
D	678	ACCATCATCTACCTCAGCTGGAGGATTTCCGCTCAGTGCACCAAGAGAACAGTCAACCGTA	737
Q	826	AAGCAACTGTTAAACACAGAGTGTGAGACCTGTGCGCGCTTTGTGAGAGAGCGCAAGGCG	885
D	738	AAGGAAAGCTGAGCGCAAGATTTACACGTTGTATGTCATCTGTGATGAGAAAGAAAGT	797
Q	886	CAACTGCTTCAACACACTGTGCCCGGAGACGAGAGGAAGTTGACGCGCTGCGGGGCTC	945
D	798	GAGTGTGTGACGCGGATCACGCGAGGACGAGGAAGAAAGCTTAAGCTTATCAGAGGCGCTC	857
Q	946	ATTCGCGCTGTACGGAGACACTTGGAGGGCTTCCTAAAGCTGTGTGAGTCCGCCATCCAG	1005
D	858	ATTCAGACGTATCCAGAGACAGCTGTGACAAAGTCCACAAAGCTGTGTGAAATCTCCATCCAG	917
Q	1006	TTCATGAGGAGCCGACAGTGGCTCTCACTCCAGACGGCAAAAGAGCTGATCAACAAG	1065
D	918	TCCCTGTGACGAGCTGTGGGAGCCACTTCTCTGTGACTGCCAAGCACTATCAAAAGC	977
Q	1066	GTGGGGCAATGTGAGAGTGTGAGCTGTGCAGAGACGCCGGAAGCCAGGCTATGAGAGCATG	1125
D	978	ATTGTGGAAGCTTCCAAAGGCTGCCAGTGTGGGGAAGACAGAACAGGCTTTGTAGAAGCAT	1037
Q	1126	GAGCAATTTCTGTGAGAGTGTGAGACAGCTGTGCCCAAAATGTTGCGAAACCATGACATCCAG	1185
D	1038	GACTCTTTACTTTTGATTTAGAGACACATACAGACGCCCTGTAGAGCCATTGACTTTGGG	1097
Q	1186	CCGGGCGCGCTGGGGATGAAGAGATGACACATGCTTTGTATGTGGGAGGAGGACGCAAT	1245
D	1098	ACAGATGAGGAAAGGAGAAATTCATTGAAAGAAATCAGGAAGAGAGAGAGATCCACA	1157
Q	1246	CGCGGCTGGAGAGGAGCGGCTGGAAGCTGCCAAGG	1283
D	1158	GAAAGGAAAGAGACACCAATTAAGAGCTGTGAG	1195

RESULT	9
AK056942	
LOCUS	1764 bp mRNA linear PRI 01-AUG-2002
DEFINITION	Homo sapiens CDNA FLJ32380 fis, clone SKMUS1000064, moderately similar to Mus musculus RING-finger protein MORF mRNA.
ACCESSION	AK056942
VERSION	AK056942.1 GI:16552479
KEYWORDS	Oligo capping; FIS (full insert sequence).
SOURCE	Homo sapiens skeletal muscle cDNA to mRNA, clone_1lb:SKMUS1

ORGANISM  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

REFERENCE  
1  
Ohtsuka, A., Takahashi-Fujii, A., Tanase, T., Inoue, N., Takeuchi, K.,  
Arita, M., Mushashino, K., Yuki, H., Hara, H., Sugiyama, T., Irie, R.,  
Otsuki, T., Sato, H., Wakamatsu, A., Ishii, S., Yamamoto, J., Isono, Y.,  
Kawai, H., Saito, K., Nishikawa, T., Kimura, K., Yamashita, H.,  
Matsuo, K., Nakamura, Y., Sekine, M., Kikuchi, H., Kanda, K.,  
Watanabe, M., Murakawa, K., Kanehori, K., Sugiyama, A., Kawakami, B.,  
Suzuki, Y., Sugano, S., Nagahari, K., Masuno, Y., Nagai, K. and  
Izumi, T.  
NEDO human cDNA sequencing project  
Unpublished  
2 (bases 1 to 1764)  
Izumi, T., Otsuki, T. and Sugiyama, T.  
Direct Submission  
Submitted (24-OCT-2001) Takao Izumi, Helix Research Institute,  
Genomics Laboratory, 1332-3 Yana, Kisarazu, Chiba 292-0812, Japan  
(E-mail: genomics@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)  
NEDO human cDNA sequencing project supported by Ministry of  
Economy, Trade and Industry of Japan; cDNA full insert sequencing:  
Research Association for Biotechnology (RAB) (supported by Japan  
key Technology Center etc.) 5'-3' end one pass sequencing: RAB,  
HRI, and Biotechnology Center, National Institute of Technology and  
Evaluation; clone selection for full insert sequencing: RAB and  
HRI.

FEATURES  
source  
1. 1764  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="SKM0100064"  
/issue\_type="skeletal muscle"  
/clone\_lib="SKM01"  
/note="cloning vector: PME18SFL3"  
137..1198  
/note="unnamed protein product"  
/codon\_start=1  
/protein\_id="BAB71318.1"  
/db\_xref="GI:16552480"  
/translation="MDYKSSLIODGNPMENLEKOLICPEMFTRPVVILPCQHNLC  
RKANDIFQANPYMTSGSSVMSGRFCPTCRHVIIMRHGVYGIQRMILVENII  
DIYKOCSSRLQKSGHPMKCKHEDEKINICLTCREPTKMSHVKELSKQPLIYA  
SVQGGKTELNCTISMLVANDRQVITITQLEBDLSTKYVFAISLDPGGATFL  
LTKOLIKSIVKSKQGLKTRQGFENMDFLIDALRALRIDFGTDEEEEFIL  
EEEDDEEEESTEGKEEGHQ"

BASE COUNT  
466 a 402 c 517 g 379 t

Query Match 33.7%; Score 482; DB 9; Length 1764;  
Best Local Similarity 66.6%; Pred. No. 1.5e-83;  
Matches 705; Conservative 0; Mismatches 350; Indels 3; Gaps 1;

QY 226 CTGCTAGGGGATGGCGCAACATGGAGAGACGCTCATTTGCCCATCTGC 285  
DB 155 CTGATTCAGAGATGGGAATCCATGAGAGACTGGAGAACGACTGATCTG 214  
QY 286 CTGAGATGTTCTCCAGCCCGTGTGATCTTGCCCTGCCAACAACCTGTGCCG 345  
DB 215 CTGAGATGTTTACCAACCCAGTGTGATCTTGCCCTGCCAACAACCTGTGCCG 274  
QY 346 TGTGCCAAGACGCTTCTGAGGCTCTGATCTCTGTGCACATCCCGGGGCGCCACAAG 405  
DB 275 TGTGCCAATGATCTTCCAGGCTGCAATCCCTACTGACCAACCGGGGAGCTCAAGT 334  
QY 406 GTGCTTCAGAGAGAGCTTCCGATGCCATCTGTAGGACGAGAGCTTGTGAGCAG 465  
DB 335 TGCATGTGTGAGAGCGCTTCCGCTGCCACACTGCCGACAGAGTGTATGATCG 394  
QY 466 CATGGTGTCTATGGCTTCAGGAGCACTGTAGTGGAGACATCATTTGACATCTCAAG 525

DB 395 CACGAGATTGAGGCTTCAGAGGAACCTGCTGTGGAGAACATCATCTACAA 454  
QY 526 CAGAGTCTCCCGCCACTGACGCGCAAGCTGAACAGCACCCTATCTGTGAGGAGC 585  
DB 455 CAGAGTCTCCCGCCACTGACGCGCGCTGCAAGAGGCGAGT - -CACCCCTGTGCAAGGAGC 511  
QY 586 GAGGCGAAGATCAACATCTACTGCTGAGCTGCGAGTGGAGTGGCCACCTGCTCTCTG 645  
DB 512 GAGATGAGAAATCAACATCTACTGCTGCTGAGCTGAGTGGAGTGGCCACCTGCTCTCTG 571  
QY 646 AAGTCTTCGCGCCGCAAGAGACTGTGAGTGGCCCTGCTGCGCCACATTTACAAAGC 705  
DB 572 AAGTCTTCGCGCCGCAAGAGACTGTGAGTGGCCCTGCTGCGCCACATTTACAAAGC 631  
QY 706 CAGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 765  
DB 632 CAAAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 691  
QY 766 GCAGTGAATCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 825  
DB 692 ACCATCATCTACAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 751  
QY 826 AAGCACTGTAAACCAAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 885  
DB 752 AAGCAAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 811  
QY 886 GAACTGCTCAAGCACTGAGGCGCGGAGGAGGAGAGAGTGAAGTGAAGTGAAGTGAAGTGAAG 945  
DB 812 GAGTGTCTCAGGCGATCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 871  
QY 946 ATCCGCGATCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1005  
DB 872 ATCCGCGATCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 931  
QY 1006 TCCATGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1065  
DB 932 TCCATGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 991  
QY 1066 GTCGCGCATGTCGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1125  
DB 992 ATTTGTGAGAGCTTCCAGAGGCTGCGAGCTGCGAGGAGGAGGAGGAGGAGGAGGAGGAG 1051  
QY 1126 GAGCAATCTCTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1185  
DB 1052 GACTTCTTACTTGTGAGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1111  
QY 1186 CCGGCGCGCGCTGCGATGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1245  
DB 1112 ACAGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1171  
QY 1246 GCGGCGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1283  
DB 1172 GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1209

RESULT 10  
AF353673 1791 bp mRNA linear PRI 25-APR-2001  
LOCUS AF353673 Homo sapiens irs ring finger protein mRNA, complete cds.  
ACCESSION AF353673  
VERSION AF353673.1 GI:13785923  
KEYWORDS  
SOURCE Homo sapiens.  
ORGANISM Homo sapiens.  
REFERENCE 1 (bases 1 to 1791)  
AUTHORS Wistow, G.  
TITLE IRF: A Novel Ring Finger Protein From Iris  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 1791)  
AUTHORS Wistow, G.

TITLE Direct Submission  
JOURNAL Submitted (27-FEB-2001) MSF, NFI, 6/331, NIH, Bethesda, MD 20892,  
USA

FEATURES  
source location/Qualifiers

1..1791  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="1"  
/map="1p33-p31"  
/tissue\_type="Iris"  
120..1181  
/note="IRP"  
/codon\_start=1  
/product="iris ring finger protein"  
/protein\_id="AAK39519.1"  
/translation="MDYKSLTODGNPMNLEKOLICPICLEMTKPVVILPCQNHLC  
RKANDIROANPYTNSRGSSYSMSGGRFCRHEVIMDRHGYGLDRLNLENI  
DIYQECSSRLQKSHPMCKRHEDEKINITYLCTEVPKSCMKVGEIHKAEVAPLO  
SVFQOKTELNNICISMLVAGNDRVQIITLQEDSRVTKENSHOVEELSHKFDALYA  
ILDERKSELQRIODEOKLSFIALIQOYEOLEDRSKVETATIOSIDEPEGATFL  
LTAOLIKSIYEASKGCOLGKTEQGFENMDFTLDELHIALDALRAIDFETDEEEEFI  
EEFOEEESTEGKEEGHQ"

BASE COUNT 505 a 398 c 512 g 376 t  
ORIGIN

Query Match 33.7%; Score 482; DB 9; Length 1791;  
Best Local Similarity 66.6%; Pred. No. 1.5e-83;

Matches 705; Conservative 0; Mismatches 350; Indels 3; Gaps 1;

226 CTGAGAGGAGTGGCCACACATGGACAACTGGAGAGACAGCTCAATTTGCCCATCTGC 285  
138 CTGATCCAGATGGGAATCCCATGGAGAACTGGAGAAACAGCTGATCTGCTATCTGC 197  
286 CTGAGAGATGTTCTCCACACCCCGTGGTGAATCTGCCTGCCACACACCTGGCCGAG 345  
198 CTGAGAGATGTTACCAACAGCAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 257  
346 TGTGCCACACAGCTGTTCCAGAGCTCTTAATCTGTGGCAATCCCGGGCTCCACAGC 405  
258 TGTGCCAATGACATCTTCCAGAGCTGCAATCCCTACTGACAGCCGCGGAGCTCAGTG 317  
406 GTGCTTCAGAGAGAGCTTTCGATGCCATCTGTAGCAGAGAGTTCCTGAGAGAG 465  
318 TCCATGCTGAGAGCCGCTTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 377  
466 CATGCTGCTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 525  
378 CACGAGAGTGTACGGCTGCAAGAGAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 437  
526 CAGAGATCTCCCGGCCACCTGACAGCCCAAGCTGAACAGCAGCTATGTGAGAGAGC 585  
438 CAGAGATGCTTCACTGAGCCGCTGCAAGAGGCACT--CACCCCATGTGCAAGAGAGC 494  
586 GAGAGAGAGAGATCAATCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 645  
495 GAGAGATGAGAAATCAATCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 554  
646 AAGTTTTCGCGCCCAAGAGAGTGTGAGTGGCCCTGCTGCTGCTGCTGCTGCTGCTG 705  
555 AAGTGTGTTGGAGTCCCAAGAGGCTGCGAGAGTGGCCCATTCAGAGAGTGTCTCAG 614  
706 CAGAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 765  
615 CAAAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 674  
766 GAGATGATCAACAGATGAGAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 825  
675 ACCATGATCACTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 734  
826 AAGCAATGTTAAACAGAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 885  
735 AAGGAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 794

886 GAATGCTTCAAGCACTGGCCCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 945  
795 GAGTGTGCTGACAGGATACCGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 854  
946 ATCCGCGAGTACGAG 1005  
855 ATCCAGAGTACGAG 914  
1006 TCCATGAG 1065  
915 TCCGAG 974  
1066 GTCCGAG 1125  
975 ATTCGAG 1034  
1126 GAGCAATCTCTGTGAG 1185  
1035 GAGTCTTACTTGTGAG 1094  
1186 CCGGAG 1245  
1095 ACAGATGAG 1154  
1246 GCGGAG 1283  
1155 GAGGAG 1192

# RESULT 11

AY059627 1861 bp mRNA linear ROD 13-DEC-2001  
LOCUS  
DEFINITION  
Rattus norvegicus muscle ring finger protein 1 mRNA, complete cds.  
ACCESSION  
AY059627  
VERSION  
AY059627.1 GI:16444655  
KEYWORDS  
Rattus norvegicus.  
Rattus norvegicus.  
SOURCE  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.  
1 (bases 1 to 1861)  
Bodine,S.C., Latre,E., Baumhueter,S., Lai,Y.K.-M., Nunez,L.,  
Clarke,B.A., Poueymirou,W.T., Panaro,F.J., Na,E., Dharmarajan,K.,  
Pan,Z.-Q., Valenzuela,D.M., Dechata,T.M., Stultz,T.N.,  
Yancopoulos,G.D. and Glass,D.J.  
Identification of ubiquitin ligases required for skeletal muscle  
atrophy  
Science 294 (5547), 1704-1708 (2001)  
JOURNAL  
MEDLINE  
21578247  
PUBMED  
11679633  
REFERENCE  
2 (bases 1 to 1861)  
Nunez,L. and Glass,D.J.  
Direct Submission  
Submitted (12-OCT-2001) Muscle Research, Regeneron Pharmaceuticals,  
777 Old Saw Mill River Road, Tarrytown, NY 10591, USA

## FEATURES

source

## CDS

1..1861  
/organism="Rattus norvegicus"  
/strain="Sprague-Dawley"  
/db\_xref="taxon:10116"  
95..1150  
/note="MurF1; monomeric E3 ubiquitin ligase"  
/codon\_start=1  
/product="muscle ring finger protein 1"  
/protein\_id="AA116405.1"  
/db\_xref="GI:16444656"  
/translation="MDYKSLTODGNPMNLEKOLICPICLEMTKPVVILPCQNHLC  
RKANDIROANPYTNSRGSSYSMSGGRFCRHEVIMDRHGYGLDRLNLENI  
DIYQECSSRLQKSHPMCKRHEDEKINITYLCTEVPKSCMKVGEIHKAEVAPLO  
SVFQOKTELNNICISMLVAGNDRVQIITLQEDSRVTKENSHOVEELSHKFDALYA  
ILDERKSELQRIODEOKLSFIALIQOYEOLEDRSKVETATIOSIDEPEGATFL  
LTAOLIKSIYEASKGCOLGKTEQGFENMDFTLDELHIALDALRAIDFETDEEEEFI  
EEFOEEESTEGKEEGHQ"



LSAKPLIKSTIVEASKGCOLGKTEGOFENMDYFTLNLHIAELRAIDFOTDEEPEETE  
 EEEDEDEGEVSTEGHO  
 BASE COUNT 490 a 433 c 534 g 404 t  
 ORIGIN

Query Match 33.3%; Score 476.8; DB 10; Length 1861;  
 Best Local Similarity 67.5%; Pred. No. 1.5e-82;  
 Matches 686; Conservative 0; Mismatches 327; Indels 3; Gaps 1;

227 TGTATAGGAGATGCGCACAACATGAGACAACTTGGAGAGACAGCTATTGGCCCATCTGCC 286  
 114 TGATTCGGAGAGGAATGATATGAGAACTGGAGAGAGCTCATCTGCCCATCTGCC 173  
 287 TGGAGATGTTCTCCAGACCCCGTGGTATCTTGGCCCTGCCAACAACCTTGGCCGAGT 346  
 174 TTGAGATGTTTACCAAGCTGTGTCTATCTGCGCCGCGCAGCAACCTTGGCCGAGT 233  
 347 GTGCCAAGAGCTTCCAGGCGCTTATCTCTGTGGCAATCCCGGGGCTCCACAACG 406  
 234 GTGCCAAGAGCATCTTCCAGGCTGCCAATCCCTACTGGAGCAACCGCGTGGCTGGT 293  
 407 TGTCTTCAAGAGAGAGCTTCCAGTCCCATCTTGTAGGCAAGAGTGTCTTGGACAGGC 466  
 294 CCATGCTGTGAGAGTGTCTTCCGCTGCGCCCTGCGCCGCGCATGAGATCATGAGACCG 353  
 467 ATGTGTCTATGAGCGCTGCGAGCGCAACCTGCTAGTGAAGCAATCATGACATGACAGC 526  
 354 ATGGGGGTAGAGGCTGTGAGAGAGCACTGTGTGAGAGCAATCATGACATCTTACAGC 413  
 527 AGGAGTCTCCCGGCGCAGCAGCGCAAGGCTGAGAGCACTCATGTGTGAGAGAGCAG 586  
 414 AGGATGCTCCAG---TGCGCCCTGCGAAGAGAGCAGCCGCGATGTGCAAGAGACG 470  
 587 AGGAGAGAGATCAACATCTACTGCTGAGCTGCGAGGTGCCACCTGCTCTCTGCA 646  
 471 AAGAGAGAGAAATCAACATCTACTGCTGAGCTGCGAGGTGCCATCTTCTTGGCA 530  
 647 AGGTTTTGGCGCCCAAGAGAGTGTGAGTGGGCGCTTGGCCACCATTTACAAAGCC 706  
 531 AGGTGTTCGGGGCTCACAGAGCTGTGAGATTGCCCCCTTACAAAGCATCTTCCAGAGC 590  
 707 AGAAGATGAGCTGAGCGAGTGCATGCGATGCTGTGGCGGGCAATGACCGTGTGAG 766  
 591 AGAAGATGAGCTGAGCGAGTGCATGCGATGCTGTGGCGGGCAATGACCGTGTGAG 770  
 767 CAGTATCACCCAGATGAGAGAGTGTGCGAGACCATTTGAGAGCAACCGCGAGACGA 826  
 651 CTATCATCTTCGAGCTGAGAGACTCTGCGAGTGCACAGAGAAACGACCAAGGTGA 710  
 827 AGCAACTGTTAAACGAGAGTGTGAGAGCTGTGCGGGTTTGGAGAGCGCAAGGCG 886  
 711 AGAGAGAGCTGAGAGCAAGTGTGAGAGCTGTGCGGGTTTGGAGAGCGCAAGGCG 770  
 887 AACTGCTTCAAGCACTGCGCGGAGAGAGAGAAAGTTGCAAGCGCTGCGGGGCTTCA 946  
 771 AGCTGCTGACAGCGAGTATCAGAGAGAGAGAAAGTGTGAGAGCGCTTCA 830  
 947 TCGCGCAGTACGAGAGACCTTGGAGAGCGCTTCAAGCTGTGAGAGCGCGCATTCAGT 1006  
 831 TCTTCCAGTACCGAGAGCTGTGAGAAAGTGTGAGAGCGCTTGTGAGAAAGCATTCAGT 890  
 1007 CCATGAGAGAGCGCAGATGCTCTTACTTCCAGAGAGAGAGAGAGCTGATCAACAAG 1066  
 891 CCCTGATGAGCGCGGAGAGCGCACTTCTCTTGTAGTCCAGAGCGCTCATCAAGAGCA 950  
 1067 TCGGGGCAATGTGAGAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1126  
 951 TTGAGAGAGCTTCCAGAGCGCTGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1010  
 1127 AGCAATCTCTGTGAGAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1186  
 1011 ACTTCTTACTGTGATTTAGAGACATAGCAGAGAGCGCTTGAAGGCGCATCTTGGGA 1070

QY 1187 CCGGCGCGCGTGGGAGATGAGAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1242  
 Db 1071 CAGATGAGAGAGAGAGAGAGTCTTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1126

RESULT 12  
 BC015717 1211 bp mRNA linear PRI 11-OCT-2001  
 LOCUS  
 DEFINITION Homo sapiens, similar to ring finger protein 28, clone MGC:17320  
 IMAGE:3922363, mRNA, complete cds.  
 ACCESSION BC015717  
 VERSION BC015717.1 GI:16041696  
 KEYWORDS MGC.  
 SOURCE Homo sapiens.  
 ORGANISM Homo sapiens.  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 1211)  
 Strausberg, R.  
 Direct Submission  
 Submitted (09-OCT-2001) National Institutes of Health, Mammalian  
 Gene Collection (MGC), Cancer Genomics Office, National Cancer  
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
 USA  
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
 Contact: MGC help desk  
 Email: [cgabbs-remail.nih.gov](mailto:cgabbs-remail.nih.gov)  
 Tissue Procurement: ATCC/DC/DT/PT  
 CDNA Library Preparation: Life Technologies, Inc.  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
 DNA Sequencing by: Genome Sequence Centre,  
 BC Cancer Agency, Vancouver, BC, Canada  
 info@bcgsc.bc.ca

REMARK  
 COMMENT  
 Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,  
 Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,  
 Leticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo  
 Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven  
 Ness, Pawan Pandoh, Anna-Liisa Pihluu, Parvaneh Saeedi, Jacqueline  
 Schein, Duane Smalhus, Michael Smith, Lorraine Spence, Jeff Stolt,  
 Michael Thorne, Miranah Tsai, Natasja van den Bosch, Jill Vardy,  
 George Yang, Scott Zuyderduyn, Marco Marra.

Clone distribution: MGC clone distribution information can be found  
 through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>  
 Series: IRAC Plate: 15 Row: O Column: 19  
 This clone was selected for full length sequencing because it  
 passed the following selection criteria: matched mRNA gi: 14211914.

#### FEATURES

##### source

1..1211  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="MGC:17320 IMAGE:3922363"  
 /tissue\_type="Skin, melanotic melanoma."  
 /clone\_lib="Nih.MGC\_72"  
 /lab\_host="DH10B"  
 /note="Vector: PCMV-SPOrt6"  
 135..1175

##### CDS

1..1211  
 /product="Similar to ring finger protein 28"  
 /protein\_id="AAH15717.1"  
 /db\_xref="GI:16041697"

1007 CCATGAGAGAGCGCAGATGCTCTTACTTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1066  
 891 CCCTGATGAGCGCGGAGAGCGCACTTCTCTTGTAGTCCAGAGCGCTCATCAAGAGCA 950  
 1067 TCGGGGCAATGTGAGAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1126  
 951 TTGAGAGAGCTTCCAGAGCGCTGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1010  
 1127 AGCAATCTCTGTGAGAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1186  
 1011 ACTTCTTACTGTGATTTAGAGACATAGCAGAGAGCGCTTGAAGGCGCATCTTGGGA 1070

BASE COUNT 336 a 302 c 345 g 228 t  
 ORIGIN  
 Query Match 33.2%; Score 475.6; DB 9; Length 1211;  
 Best Local Similarity 68.1%; Pred. No. 2.6e-82;

Matches	677;	Conservative	0;	Mismatches	314;	Indels	3;	Gaps	1;
OY	226	CTGCTAGGGGATGCGCAACATGACACTTGGAGAGAGACTGATTTGCCCATTCG							
Db	153	CTGATTCAGAGATGGGATCCATGAGAACTTGGAGAGAGAGCTGATCTGCCATCTGC							
OY	286	CTGAGAGATGTTTCCAGAGCCGCTGTGATCTTGCCTGCGCAACAACTGTGGCCGAG							
Db	213	CTGAGAGATGTTTACCAAGCAGAGCTGATCTTCCGTCGACAGCAACACTGTCCGAG							
OY	346	TGTGCCAAGAGACTCTTCCAGGCTTAAATCTCTGTGCAATCCGGGGCTCCACAGC							
Db	273	TGTGCAATGATATCTTCCAGGCTGCAAAATCCCTACTGACACAGCGGGGACACTGAGT							
OY	406	GTGTCTAGAGAGAGAGCTTCCGATGCCATCTGTAGAGAGAGAGTGTCCGAGACAG							
Db	333	TCCATGTCTGAGAGGCGCTTTCCGCTGCCACCTGCGCCGACAGGATGATGAGATCGT							
OY	466	CATGATGCTATGAGCGGAGGAACTGTAGTGGAGAGACATCATTTGATCTACAG							
Db	393	CACGAGTGTACGCGCTGAGAGAACTGTGTGGAGAACATCATTCGACATCTACAA							
OY	526	CAGAGTCTCTCCCGGCTGACCCGCAAGGCTGAACACACACTCATGTGTGAGAGAC							
Db	453	CAGGAGTCTCCAGTGGGCGCTGACAGAG---GGCAGTACCCCACTGTGCAAGAGAC							
OY	586	GAGAGCAGAGATCAATCATCTACTGCTGAGCTGGAGAGTGGCCGCTGCTCTCTGC							
Db	510	GAAGATGAGAAATCAATCATCTACTGCTGAGCTGGAGAGTGGCCGCTGCTCACTGTC							
OY	646	AAAGTTTTCGCGCCCAAGAGACTGTGAGTGGCCCTCTCCCAACATTTACAAACGC							
Db	570	AAAGTGTGTGGATCCCAAGAGGCTGAGAGTGGCCCTCTCCAGAGTGTCTTCCAGGGA							
OY	706	CAGAGAGTGTGAGAGGATGAGTGGATGCGATGCTGTGTGGCCGCAATACCGTGTGAG							
Db	630	CAAAAGACTGACTGATTAATCACTGATCTCATGCTGTGTGGGGGAGATGACCGTGTGAG							
OY	766	GCACTGATCACCCAGATGAGAGAGTGTGCCAGACATTGAGAGACAAAGCCGACAGAC							
Db	690	ACCATCTACATCTGAGAGATCCCGTGAATGACCAAGAGAACAGTCAACCGATG							
OY	826	AAGCACTGTTAAACAGAGTGTGAGACCTGTGTGCGGCTTTGGAGAGCGCAAGGCG							
Db	750	AAGCAAGAGCTGAGCCAGAAATTTGACAGCTGTGATGCACTCTGATGAGAAAGAACT							
OY	886	GAATGCTTCAAGCACTGCGCCGAGAGAGAGAGAGATTGACGCGGTGCGGGCGCTC							
Db	810	GAGTTGCTGAGCGGATTCACGACAGAGAGAGAGAAAGCTTACCTTATCGAGGCGCTC							
OY	946	ATCGCCAGTACGAGACCACTTGGAGGGCTCTCAAAAGCTGTGTGAGTCCGCCATCCAG							
Db	870	ATCCAGAGATACCAAGGAGACAGCTGAGCAAGTCCACAAAGCTGTGTGAAACCTCCATCCAG							
OY	1006	TCCATGAGAGAGCGCGAGATGCTCTCTACTCCAGCAGAGCAAGAGACTGATCAACAAG							
Db	930	TCCCTGACAGAGCTTGGGGAGGACCACTTCTCTTGTGACTGCGCAACCACTCAAAAGC							
OY	1066	GTGCGGGCAATGTCGAGAGTGGAGCGAGAGCGCGGAGCCAGCCTATGAGAGAGT							
Db	990	ATTGTGAAAGCTTCCAAAGGGGTGCCAGTGGGGAAGACAGAGAGGCTTTGAGAACATG							
OY	1126	GAGCAATTTCTGTGAGCGTGGAGACAGTGGCCGAATGTTGCAACCACTGACTTCAG							
Db	1050	GACTTCTTACTTTGATTTAGAGACACATAGCAGAGCGCTGAGACCTATGACTTTGGG							
OY	1186	CCGGGCGCGCTGGGATGAAAGAGATGACACA							
Db	1110	GCAAGCTGAAGATGATGAGATGAGATGTGATA							

RESULT 13  
AX418850

LOCUS	AX418850	2590 bp	DNA	linear	PAT 18-JUN-2002
DEFINITION	Sequence 3 from Patent WO0206318.				
ACCESSION	AX418850				
VERSION	AX418850.1	GI:21523714			
KEYWORDS					
SOURCE	house mouse.				
ORGANISM	Mus musculus.				
REFERENCE	Olson, E.N. and Spencer, J.A.				
AUTHORS	Methods and compositions for stabilizing microtubules and				
TITLE	Intermediate filaments in striated muscle cells				
JOURNAL	Patent: WO 0206318-A 3 24-JAN-2002;				
FEATURES	BOARD OF REGENTS THE UNIVERSITY OF TEXAS SYSTEM (US)				
source	1. .2590				
CDS	/organism="Mus musculus"				
	/db_xref="taxon:10090"				
	80..1717				
	/note="unnamed protein product"				
	/codon_start=1				
	/protein_id="CAD35456.1"				
	/db_xref="GI:21523715"				
	/translation="MSTSLATKSKSKQQTMDNLEKILICPLEMTKPYVILPCOH				
	NLCRKASDILFQASNPFLPTRGTTVASGGRFCPSRHEVYLDHRHGVYGLRNILYE				
	NIIIDIKQESTREKRLIDPMECEHEBERINILYCLNEVPTCSLCKEYVAGRKDOVAP				
	LTHFQOKSELDGIYLVGSDNRVQYISOLDPTEKTECCRKQDKODCEKEDHL				
	YGLLEERKTEMTQALPTROEKELEHVEITLRKYSDHLENVSKLVEGLOPDEBMY				
	FLNARTLLOKIVASKAFQMEKLEQGYELINMNTVYINPREKILIREIDPSREEEE				
	DAEIEDEEGEDVAEVEEAPENNOIASGSESESELEKARPSQLFPAELQVAPRLPAS				
	PEPSSMPPRADIVYDGEVYPLGSQDTTQSEISGPAAEADPLFPMWYKQSRKT				
	SNMPCITHEGEGVQIQLPDIEDSVQSAVEAEATNEQAAVSGKSSSTPATSOIGP				
	EAPSPGQSALISGGGVILSOLATSPSPGLNSLND"				
BASE COUNT	703	a	627	c	625
ORIGIN	635 t				
Query Match	31.5%; Score 450.8; DB 6; Length 2590;				
Best Local Similarity	63.8%; Pred. No. 1.6e-77;				
Matches 702; Conservative	0; Mismatches 392; Indels 6; Gaps 1;				
OY	187	GAGCGCGCGGAGTGAACCTTACGCGTGTTCAGCGCGCTGATGAGGATGCGCACAAAC			
Db	68	GAGCGCAAGAAATGACACACTTCTCTGATTAACAAGTCTTCTCCAAAGACAGACAGCC			
OY	247	ATGGAACACTTGGAGAGAGAGCTCATTTGCCCATCTGCGCTGAGAGTGTCTCCAAGCC			
Db	128	ATGATTAACCTTGGAGAGAGAGCACTGATCTGCCATCTGCGCTGAGAGATGTCACGAAGCT			
OY	307	GTGATGATCTTGCCTGCGCAACACCTGAGCGGCAAGGTGTCACAGAGAGCTTCCAG			
Db	188	GTGCTCATTTCTCCCTTGCACAGCAACCTGTGCAAGAAATGTGCAAGTATCTCCAG			
OY	367	GCTTAAATCTCTGTGGCAATCCCGGGGCTCCACAAAGGTGTTTGAAGAGAGAGCTTC			
Db	248	GCTTAACCCGTAATCTTACCAAGAGAGAGCAACCGGGGATCCAGGAGGCGGCTTC			
OY	427	CGATGCCAATCTTGTAGGCGAGAGTGTCTCGACAGGATGTGTCTATGGCTCGAG			
Db	308	CGCTGTCCTCTGCGAGACTGAGGTGTGTAGACGAGATGGGCTCTATGAGCTCGAG			
OY	487	CGAATCTGATGAGAGACATCATCTACAGAGAGAGAGTCTCCCGGCGCACTG			
Db	368	AGGAACCTGTGTGGAAGAAATATTGATATCTACAGAGAGAAATCCACAGGCGCAA			
OY	547	CACGCCAAGGCTGACAGACACTCATGTGTGAGAGCAGAGAGAGAGAAATCAATCATC			
Db	428	AAAAAATTTG-----ACGAGCCCATGTGTGAAGAGCATGAAGAAACCATCAATCATC			
OY	607	TACTGCTGAGGTGCGAGGAGCGCCACCTGCTCTGCAAGGTTTTCGCGCCGCAAG			
Db	482	TATTTGCTAACTGTGAAGTGGCCACCTGTTCTGTGCAAGGTTTTCGCGCCATTAAG			



```

QY 667 GAGCTGTGAGGTGGCCCTCTGCGCCACATTACAAAGCCAGAGAGATGAGTGCAT 726
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 542 GAGCTGTGAGGTGGCCCTCTGCGCCACATTACAAAGCCAGAGAGATGAGTGCAT 601
QY 727 GGCATCCGCGATGCTGTGGCGGCAATGACCGCTGTGCGAGGAGTATCACCAGATGAG 786
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 602 GGTATGCTGTACTGTGTGGAGGACAAAGATGATGATGATGATGATGATGATGATGAT 661
QY 787 GAGGTGTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 846
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 662 GAGACCTGTAAACATTAATGATGATGATGATGATGATGATGATGATGATGATGAT 721
QY 847 TTGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 906
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 722 TTGATGACCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 781
QY 907 CGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 966
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 782 CGAGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 841
QY 967 TTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1026
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 842 CTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 901
QY 1027 GCTCTCTACCTCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1086
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 902 GAGATATTTCTGAGAGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 961
QY 1087 GAGCTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1146
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 962 TTTCAGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1021
QY 1147 GAGCAGCTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1206
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1022 AATGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1081
QY 1207 GAGGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1266
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1082 GATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1141
QY 1267 CTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1286
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1142 GAAATGATTCATTAATGATC 1161

```

RESULT 14  
 BC007750 1750 bp mRNA linear PRI 12-JUL-2001  
 LOCUS BC007750  
 DEFINITION Homo sapiens, clone MGC:12836 IMAGE:4110783, mRNA, complete cds.  
 ACCESSION BC007750  
 VERSION BC007750.1 GI:14043531  
 KEYWORDS MGC.  
 SOURCE Homo sapiens.  
 ORGANISM Homo sapiens.  
 Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
 1 (bases 1 to 1750)  
 Strausberg, R.  
 Direct Submission  
 Submitted (11-MAY-2001) National Institutes of Health, Mammalian  
 Gene Collection (MGC), Cancer Genomics Office, National Cancer  
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
 USA  
 NIH-MGC Project URL: <http://mgc.ncl.nih.gov>  
 Contact: MGC help desk  
 Email: [cgabs-remail.nih.gov](mailto:cgabs-remail.nih.gov)  
 Tissue Procurement: ATCC  
 cDNA Library Preparation: Rubin Laboratory  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILIN)  
 DNA Sequencing by: National Institutes of Health Intramural  
 Sequencing Center (NISC),  
 Gaithersburg, Maryland:  
 Web site: <http://www.nisc.nih.gov/>

Contact: [nisc\\_mgc@nhgri.nih.gov](mailto:nisc_mgc@nhgri.nih.gov)  
 Shevchenko, Y., Wetherby, R.D., Beckstrom-Sternberg, S.M.,  
 Benjamin, B., Blakesley, R.W., Bouffard, G.G., Brinkley, C., Brooks, S.,  
 Dietrich, N.L., Guan, X., Gupta, J., Ho, S.-L., Karlins, E., Legaspi, R.,  
 Lim, M., Maduro, O.L., Masfello, C., Mastrian, S.D., McCloskey, J.C.,  
 McDowell, J., Pearson, R., Snyder, B., Stantirpop, S., Thomas, P.J.,  
 Tlionson, E.E., Touchman, J.W., Tsurgon, C., Vogt, J.L., Walker, M.A.,  
 Zhang, L.-H. and Green, E.D.

Clone distribution information can be found  
 through the I.M.A.G.E. Consortium/ILIN at: <http://image.llnl.gov>  
 Series: IRAL Plate: 17 Row: 6 Column: 4  
 This clone was selected for full length sequencing because it  
 passed the following selection criteria: Hexamer frequency ORF  
 analysis, Similarity but not identity to protein.

Location/Qualifiers  
 1. 1750  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="MGC:12836 IMAGE:4110783"  
 /tissue="Muscle, Rhabdomyosarcoma"  
 /clone\_id="NIH\_MGC\_17"  
 /lab\_host="DH10B-R"  
 /note="vector: pOTB7"  
 190..1500  
 /codon\_start=1  
 /product="Unknown (protein for MGC:12836)"  
 /protein\_id="AAH07750.1"  
 /db\_xref="GI:14043532"  
 /translation="MDNLEKOLICPLEMTPVYLPCOHLGRKASDIFQASNP  
 LTPFRGTTMASGGRFRRPSCHREYVLDHRVYGLORNLININIDIVYQESTRPERK  
 SDPQCEHEERINITYCLNCEVPTCSLDKVGARHDCOVAPLHFRQKSLSTSGI  
 AILVGSNDRVQVLSQIDPCTKTEECCKRQKQELCEKDYLYGLIEERNEVYIT  
 RTQEKLEHVALIKKYSDBLENVKLYESGIQFMDPEMAVFLQAKTLKLSIAS  
 KAFQMEKLEHGVNHNFTVNLNREKILREIDFYEDDEDEEGBEGKEGEGVG  
 GAIVEVEVENYQTEPGENDEPKASELIREIDFYEDDEDEEGBEGKEGEGVG  
 PVTQIGFEPPLQGGAAAPASSGSSGSEARHIFSRMSLSNLSNE"  
 ORIGIN  
 BASE COUNT 556 a 389 c 433 g 372 t  
 Query Match 31.5%; Score 450.6; DB 9; Length 1750;  
 Best Local Similarity 64.8%; Pred. No. 1.8e-77;  
 Matches 703; Conservative 0; Mismatches 374; Indels 8; Gaps 2;

```

QY 187 GAGCGCGCGGGGATGAGCTTACGGTGGTTCAAGCCGCTGCTAGGGGATGCCCAAC 246
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 130 GACAGCGAGAGATGAGCGCATCTGCAATTAACAAATCTTTCCAAAGAGACAGAC 189
QY 247 ATGGAACAATCTGAGAGAGAGAGCTATTTGCCCATCGCTGGAGATGTTCCAGGCC 306
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 190 ATGGAATTAATGAGAGAGAGAGCTATCTGCTGCTTACAGATGTTTACGAACCT 249
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 307 GTGATGATCTTGGCTGCGCAACACAACTTGGCGCAAGTGTGCCAAGAGCTTCCAG 366
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 250 GTGCTGATTTCTCCCTGTGACGACAACTGTGTGAGAAATGTGCAGATATTTCCAG 309
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 367 GCCTTAATCTCTGTGCGCAATCCCGGGCTCCACAAAGGTTCTTACAGAGAGCTTC 426
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 310 GCCTTAATCCCTGATTTGCCCAAGAGAGAGGTACCAACATGGATATGAGGGCGGATTC 369
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 427 CGATGGCATCTTGTAGACAGAGGTTGCTCGACAGAGAGTGTGTTAGTCTGAGCTG 486
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 370 CGCTGCCATCTGTATACATGAAGTGTGTTGATATGACATGGGATATGAGCTTAC 429
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 487 CGGAACCTGTAGTGAAGAACATCATTGACATCTTACAGAGAGAGTCTCCCGGCACTG 546
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 430 AGGAACCTGTGTTGGAAATATATCATATGACATCTTACAGAGAGAGTCCACAGGCCA 486
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 547 CAGGCCAAGCTGAGAGAGACATCTATGTGTAGAGAGAGAGAGAGAGAGAGAGAG 606
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 487 ---GAAAAGAAATCGACAGCCAGCTGTGCGAGAGAGAGAGAGAGAGAGAGAG 543
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 607 TACTGCTGAGCTGCGAGGTGCCACCTGCTCTCTGCAAGGTTTTCGGGCGCCCAAG 666

```

```

Db 544 TACTGCTCTAACTGCGAAGTACCCACCTGCTCTGTGCGAAGGTGTTGGTGCACACAA 603
Oy 667 GACTGAGAGTGGGCCCCCTTGGCCACCATTTACAAGCCGAGAGAGTGAAGCCAT 726
Db 604 GACTGCTGAGTGGTCCCTCCCTCTCACTGTGTTCAGAGAGAGAGTCTGAGCTCAGTAT 663
Oy 727 GGCATCGGATCGTGGTGGCGGCGATGAGCCGCTGTCAGAGCAGTGAATCACCAGATGAG 786
Db 664 GGCATCGGATCGTGGTGGCGGCGATGAGCCGCTGTCAGAGCAGTGAATCACCAGATGAG 723
Oy 787 GAGGTGCGCAGACCATTTAGAGCAACGCCGAGAGAGAGTGTAAACAGAG 846
Db 724 GACACCTCTCAAACTATCTGAGATGTTGAGAAACAGAAACAGAGCTTTGTGAGAG 783
Oy 847 TTGAGACCCCTGTGGCGGGTTTTGGAGAGCCGAGGCGAATGCTCTGACAGCAGTGGCC 906
Db 784 TTTGATTTACCTGTATGAGCATTTTGGAGAGAGAGAGAAATGAAATGACCAAGTCATTACC 843
Oy 907 CGGAGCAGAGAGAGAGAGTTCGAGCGCGTGGCGGCTCATCGGCGATGAGAGACAC 966
Db 844 CGAAGCCCAAGAGAGAGAACTGGAACATGTCGCTGCTGATCAAAAGTATTTGATCAT 903
Oy 967 TTGAGGGCTCTCTCAAACTGCTGAGATGTCGCCATCCATGAGAGAGAGCCGAGAGT 1026
Db 904 TTGGAGAACGTTCTCAAACTGTTGTTGAGTCAAGAAATTCAGTTTATGATGAGCCAGAAATG 963
Oy 1027 GCTCTCTACTCTCAAGCAGAGCAAGAGAGTATCAACAAAGTGGGCGAATGTCAGAGT 1086
Db 964 GCAAGTGTTCGAGAGAAATGCCAAACCTGCTATMAAAATATTCGAGAACATCAAGAGCA 1023
Oy 1087 GAGCTGCGAGAGAGCGCGGAGAGCGATATGAGAGCAGTGAATTCCTGTGAGCGT 1146
Db 1024 TTTGAGATGAGAAATATGAGATGCTATGAGAACATGAGAACACCTTACAGATCAACCT 1083
Oy 1147 GAGCAGCTGGCGGAAATGTTGGAACCATGCACTCCAGCCCGGCGGCTGGGGATGAA 1206
Db 1084 AATGAGAGAGAAAGATATATACGTAAATTTGACITTTACAGAGAAAGATAAATGAGAA 1143
Oy 1207 GAGGATGACGATGCTTTGATGAGAGAGAGCAATGCGGGCTGAGAGAGAGAGCG 1266
Db 1144 GAGAGAGAGAGAGCG--GAGAGAGAGAGAGAAAGAGAGAGAGAGAGAGAGAGAG 1201
Oy 1267 CTGGA 1271
Db 1202 TAGAA 1206

RESULT 15
AK091728 1925 bp mRNA linear PRI 15-JUL-2002
LOCUS AK091728
DEFINITION Homo sapiens cDNA FLJ34409 fls, clone HEART2001931, moderately
ACCESSION AK091728.1 GI:21750167
VERSION AK091728.1 GI:21750167
KEYWORDS Oligo capping; fls (full insert sequence).
SOURCE Homo sapiens heart cDNA to mRNA, clone lib:HEART2
clone:HEART2001931.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 Nishi, T., Ota, T., Nakagawa, S., Senoh, A., Mizuguchi, H., Inagaki, H.,
Sugiyama, T., Ito, R., Otsuki, T., Sato, H., Wakamatsu, A., Ishii, S.,
Yamamoto, J., Isono, Y., Kawai-Hio, Y., Saito, K., Nishikawa, T.,
Kimura, K., Yamashita, H., Matsuo, K., Nakamura, Y., Sekine, M.,
Kikuchi, H., Kanda, K., Magatsuma, M., Murakawa, K., Kanehori, K.,
Takahashi-Fujii, A., Oshima, A., Sugiyama, A., Kawahara, B., Suzuki, Y.,
Sugano, S., Nishihara, K., Masuno, Y., Nagai, K. and Isogai, T.,
NEDO human cDNA sequencing project
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1925)
AUTHORS Isogai, T. and Yamamoto, J.

```

```

TITLE Direct Submission
JOURNAL Submitted (04-JUL-2002) Takao Isogai, FLJ Project(HRI Team); 2-6-7
Kazusa-Kamatari, Kisarazu, Chiba 292-0812, Japan
(E-mail: genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
NEDO human cDNA sequencing project supported by Ministry of
Economy, Trade and Industry of Japan; cDNA full insert sequencing:
Research Association for Biotechnology (RAB); cDNA library
construction: Helix Research Institute (HRI) (supported by Japan
Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB,
HRI, and Biotechnology Center, National Institute of Technology and
Evaluation; clone selection for full insert sequencing: HRI and
RAB; annotation: HRI and RAB.
FEATURES
source
1..1925
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="HEART2001931"
/tissue="heart"
/clone_lib="HEART2"
/note="cloning vector: PME18SFL3"
BASE COUNT 579 a 441 c 498 g 407 t
ORIGIN
Query Match 31.5%; Score 450.6; DB: 9; Length 1925;
Best Local Similarity 65.4%; Pred. No. 1.8e-77;
Matches 675; Conservative 0; Mismatches 351; Indels 6; Gaps 1;
Oy 187 GAGCGCGCGGAGTAACCTGACGCTGTTCAAGCCGCTGTAAGGATGCGCAAC 246
Db 35 GAGCAGAGAGAGAGAGAGCGATCTCTGTAATTAACAATCTTTTCAAGAGAGAGAGAC 94
Oy 247 ATGAGACACTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 306
Db 95 ATGATATCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 154
Oy 307 GTGATGATCTTGGCTGCGCAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 366
Db 155 GTGATGATCTTGGCTGCGCAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 214
Oy 367 GCTCTATCTCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 426
Db 215 GCTCTATCTCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 274
Oy 427 CGATGCCATCTTGTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 486
Db 275 CGCTGCCATCTTGTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 334
Oy 487 CGGACCTGCTAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 546
Db 335 AGGACCTGCTGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 391
Oy 547 CAGCGCAAGGCTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 606
Db 392 ---GAAAGAAATCCGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 448
Oy 607 TACTGCTGAGTGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 666
Db 449 TACTGCTGAGTGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 508
Oy 667 GACTGAGAGTGGCCCCCTGCCCCACCATTTACAAAGCAGAGAGAGAGAGAGAGAGAT 726
Db 509 GACTGAGAGTGGCCCCCTGCCCCACCATTTACAAAGCAGAGAGAGAGAGAGAGAGAT 568
Oy 727 GGCATGCGATGCTGTGGCGGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 786
Db 569 GGCATGCGATGCTGTGGCGGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 628
Oy 787 GAGGTGCGCAGACCATTTAGAGCAACGCCGAGAGAGAGAGAGAGAGAGAGAGAGAG 846
Db 629 GACACCTCTCAAACTATCTGAGATGTTGAGAAACAGAAACAGAGCTTTGTGAGAG 888
Oy 847 TTGAGAGCCCTGTGGCGGGTTTTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 906

```

Db 689 TTGATACCTGATGCGATTTTGGAGGAGGAGGAATGAATGACCAAGCTCATTTAC 748  
 QY 907 CGGAGCAGAGAGAGAGAGTTCAGCGCGTGGCGGCTCATCCGCAATGACGAGACCAC 966  
 Db 749 CGAACCCAGAGAGAGAGAGTTCAGCGCGTGGCGGCTCATCCGCAATGACGAGACCAC 808  
 QY 967 TTGGAGGCGCTCTCAACAGTGTGGAGTGGCGGCTCATCCGCAATGAGGAGCGCGAGAT 1026  
 Db 809 TTGGAGGAGGCTCTCAACAGTGTGGAGTGGCGGCTCATCCGCAATGAGGAGCGCGAGAT 868  
 QY 1027 GCTCTCTACTCTCCAGCAGCAAGAGAGTATCAACAGTGGCGGCTCATCCGCAATGAGGAG 1086  
 Db 869 GCAAGTGTCTTTCAGCAAGAGAGAGTATCAACAGTGGCGGCTCATCCGCAATGAGGAG 928  
 QY 1087 GAGCTGCGAGAGAGAGAGAGTATGAGAGAGTATGAGAGAGTATGAGAGAGTATGAGAG 1146  
 Db 929 TTTTCAGATGAGAGAGAGAGAGTATGAGAGAGTATGAGAGAGTATGAGAGAGTATGAGAG 988  
 QY 1147 GAGCAGCTGGCGGAGAGAGTATGAGAGAGTATGAGAGAGTATGAGAGAGTATGAGAGAG 1206  
 Db 989 AATGAGAGAGAGAGAGAGTATGAGAGAGTATGAGAGAGTATGAGAGAGTATGAGAGAGAG 1048  
 QY 1207 GAGGATGAGAGAG 1218  
 Db 1049 GAGGATGAGAGAG 1060

RESULT 16  
 HSA291712 2202 bp mRNA linear PRI 06-JUL-2001  
 DEFINITION Homo sapiens mRNA for RNF29 gene for ring finger protein 29.  
 ACCESSION AJ291712.1 GI:13160384  
 VERSION 1  
 KEYWORDS ring finger protein 29; RNF29 gene.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 2202)  
 Centner, T., Yano, J., Kimura, E., McElhinny, A.S., Pellu, K., Wilt, C.C., Bang, M.L., Trombitas, K., Granzier, H., Gregorio, C.C., Sorimachi, H. and Labail, S.  
 Identification of muscle specific ring finger proteins as potential regulators of the titin kinase domain  
 J. Mol. Biol. 306 (4), 717-726 (2001)  
 JOURNAL  
 MEDLINE 21140140  
 PUBMED 11243782  
 2 (bases 1 to 2202)  
 Centner, T.  
 Direct Submission  
 Submitted (12-FEB-2001) Centner T., Structure and Biocomputing, EMBL-Heidelberg, Meyerhofstrasse 1, Heidelberg 69112, Germany  
 JOURNAL  
 REMARK Revised by author 22-FEB-2001  
 FEATURES  
 location/Qualifiers  
 1..2202  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 133..135  
 /note="putative alternative translation initiation codon"  
 /evidence="not\_experimental"  
 181..1779  
 /gene="RNF29"  
 181..1779  
 /note="RNF29"  
 /note="alternative"  
 /codon\_start=1  
 /product="ring finger protein 29"  
 /protein\_id="CAC32840.1"  
 /db\_xref="GI:13160384"  
 /translation="MDNLEKOLIPICLEMFTKPVLLPCOHNLCRKASDIFQASNP  
 YLPRTGRTMASGRFCPCSRHEVYLDHRVYGLORNLVENIIDIKOESTREBK  
 SDPMCEHEERINICLNCNEVPTSCSKFGAHKQOAVPLTHVFOROKSEISDGI  
 AIIVGSNDRVQVITSLQLEDYTCITIECCRKOKELCEKFDYLLGLERKKNEMTOYIT  
 AIIVGSNDRVQVITSLQLEDYTCITIECCRKOKELCEKFDYLLGLERKKNEMTOYIT

CDS  
 join(181..1365,1654..1779)  
 /gene="RNF29"  
 /note="short isoform"  
 /codon\_start=1  
 /product="ring finger protein 29"  
 /protein\_id="CAC32839.1"  
 /db\_xref="GI:13160385"  
 /translation="MDNLEKOLIPICLEMFTKPVLLPCOHNLCRKASDIFQASNP  
 YLPRTGRTMASGRFCPCSRHEVYLDHRVYGLORNLVENIIDIKOESTREBK  
 SDPMCEHEERINICLNCNEVPTSCSKFGAHKQOAVPLTHVFOROKSEISDGI  
 AIIVGSNDRVQVITSLQLEDYTCITIECCRKOKELCEKFDYLLGLERKKNEMTOYIT  
 RTOBEKLEHVALIKKYSDDLHNYSKLVESGIQFDEBMAFIONATLTKISEAS  
 KAFOMEKIEHGYEMNNHFTVNLNREKILIREIDYREDEEBSGEGBEKGEVIG  
 GAFAVEVEVENVOETEPGDEBNPEKASELSOVELQAAAGALVPSPEPPALPADA  
 PVTOIGFAPPLQSOAAPASGSDSEPARHIFSFWLNSLNE"

BASE COUNT 660 a 489 c 560 g 493 t  
 ORIGIN  
 Query Match 31.5%; Score 450.6; DB 9; Length 2202;  
 Best Local Similarity 64.8%; Pred. No. 1.8e-77;  
 Matches 703; Conservative 0; Mismatches 374; Indels 8; Gaps 2;  
 QY 187 GAGCCCGCGGATGACCTTCAGCGGTTCACCGCGTGTAGGGATGCGCACAC 246  
 Db 121 GACACCGAGAGATAGCCCATCTGAAATACAAATCTTTCCAAAGACGACGACC 180  
 QY 247 ATGACAACTTGAGAGAGAGAGTTCGCCATCTGCGTGGAGATGTTCCAAAGCC 306  
 Db 181 ATGATTAATTTAGAGAGAGAGAGTTCGCCATCTGCGTGGAGATGTTCCAAAGCC 240  
 QY 307 GTGTGATCTTGGCTGCCAACAACCTGTGCGGAGTGTGCCAAGCTTCCAG 366  
 Db 241 GTGTGATCTTGGCTGCCAACAACCTGTGCGGAGTGTGCCAAGCTTCCAG 300  
 QY 367 GCCTTAATCTCTGTGGGAATCCGGGCTCCACAAGGTGTCTTCCAGAGAGAGCTTTC 426  
 Db 301 GCCTTAATCTCTGTGGGAATCCGGGCTCCACAAGGTGTCTTCCAGAGAGAGCTTTC 360  
 QY 427 CGATGCCATCTTGGAGAGAGAGTTCGAGAGAGAGAGTGTGATGAGCTTCCAG 486  
 Db 361 CGATGCCATCTTGGAGAGAGAGTTCGAGAGAGAGAGTGTGATGAGCTTCCAG 420  
 QY 487 CGAAGCTGTAGTGAAGAGAGAGTTCGAGAGAGAGAGTGTGATGAGCTTCCAG 546  
 Db 421 AGGAACTGTGTGAGAGAGAGTTCGAGAGAGAGAGTGTGATGAGCTTCCAG 477  
 QY 547 CAGCCAGAGGTGAGAGAGAGTTCGAGAGAGAGAGTGTGATGAGCTTCCAG 606  
 Db 478 ---GAAAAGAAATCCGACAGAGAGTTCGAGAGAGAGAGTGTGATGAGCTTCCAG 534  
 QY 607 TACTGTGAGAGTGAAGAGAGTTCGAGAGAGAGAGTGTGATGAGCTTCCAG 666  
 Db 535 TACTGTGAGAGTGAAGAGAGTTCGAGAGAGAGAGTGTGATGAGCTTCCAG 594  
 QY 667 GACTGTGAGAGTGGCCCTTCCGACCATTTACAAAGCCGAGAGAGAGTGTGATGAGCTTCCAG 726  
 Db 595 GACTGTGAGAGTGGCCCTTCCGACCATTTACAAAGCCGAGAGAGAGTGTGATGAGCTTCCAG 654  
 QY 727 GGCATGCGAGTGTGAGAGAGAGTTCGAGAGAGAGAGTGTGATGAGCTTCCAG 786  
 Db 655 GGCATGCGAGTGTGAGAGAGAGTTCGAGAGAGAGAGTGTGATGAGCTTCCAG 714  
 QY 787 GAGGTGTGAGAGAGAGTTCGAGAGAGAGAGTTCGAGAGAGAGTGTGATGAGCTTCCAG 846  
 Db 715 GAGGTGTGAGAGAGAGTTCGAGAGAGAGAGTTCGAGAGAGAGTGTGATGAGCTTCCAG 774  
 QY 847 TTGAGAGAGTGTGAGAGAGAGTTCGAGAGAGAGAGTTCGAGAGAGAGTGTGATGAGCTTCCAG 906

Db 775 TTTGATTACCTGATGATGCAATTTTGGAGAGAGAGAAATGACCAAGTCAATTAAC 834  
 Oy 907 CGGAGACAGAGAGAGAGAGATTTGACAGCCGCTGCGGCTCTATCCGCCAGTACGAGACCAC 966  
 Db 835 CCAACCCAAAGAGAGAGAACTGGAACATGTCCTGCTGATCAAAAAGTATTCTGATCAT 894  
 Oy 967 TTGGAGAGGCTCTCAAAAGCTGTGAGTCGCCATCCATGCAATCGATGGAGCCGACATAT 1026  
 Db 895 TTGGAGAACGTCCTCAAAAGTTGGTTGATGATGAGAAATTCAGTTTATGATGAGCCAGAAATG 954  
 Oy 1027 GCTCTACTCTCCAGCAGCAGCAAGAGCTGATCAACAGAGTCGGGGCAATGTCGAAGTG 1086  
 Db 955 GCAGTGTCTCTCAGATATCCCAAAACCTGCTGTAATAAAATTCGGAAGCATCAAGACGA 1014  
 Oy 1087 GAGTCGGCAGAGCGCGGAGCAGCAGTATGAGAGATGAGAGCAATTTCTGTGAGCGTG 1146  
 Db 1015 TTTTCAGATGAGAGAAATAGAACATGCTATGAGAACATGAAACACCTTCACAGTCAACCTC 1074  
 Oy 1147 GAGCAGTGGCCGAAATGTTGGCAACATGACATTCAGCCGCGCGCTGGGGATGAA 1206  
 Db 1075 AATAGAGAAAGAAAGATATAGCTGAATTTGACTTTTACAGAGAAATGAAATGAAAGA 1134  
 Oy 1207 GAGCATGACAGATGCTTTGATGAGGCGAGAGGCAATGCGGGCTGAGAGAGAGCGG 1266  
 Db 1135 GAGAGAGAGAGC--GGAG 1192  
 Oy 1267 CTGGA 1271  
 Db 1193 TAGAA 1197

RESULT 17  
 AK091310  
 LOCUS 2634 bp mRNA linear PRI 15-JUL-2002  
 DEFINITION Homo sapiens cDNA FLJ3991 fls, clone DNE52007332, moderately similar to Mus musculus RING-finger protein MORF mRNA.  
 ACCESSION AK091310  
 VERSION AK091310.1 GI:21749650  
 KEYWORDS oligo capping; fls (full insert sequence).  
 SOURCE Homo sapiens normal dermal fibroblasts (Neonatal Skin) (NHDF2564)  
 ORGANISM Homo sapiens  
 HOMO sapiens clone\_1lb:DNE52 clone:DNE52007332.  
 Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
 REFERENCE 1  
 Tashiro, H., Yamazaki, M., Watanabe, K., Kumagai, A., Itakura, S., Fukuzumi, Y., Fujimori, Y., Komiyama, M., Sugiyama, T., Irie, R., Otsuki, T., Sato, H., Wakamatsu, A., Ishii, S., Yamamoto, J., Isono, Y., Kawai-Hio, Y., Saito, K., Nishikawa, T., Kimura, K., Yamashita, H., Matsuo, K., Nakamura, Y., Sekine, M., Kikuchi, H., Kanda, K., Waga, S., Mura, M., Murakawa, K., Kanehori, K., Takahashi, F., Oshima, A., Sugiyama, A., Kawakami, B., Suzuki, Y., Sugano, S., Nagahara, K., Masuh, Y., Nagai, K. and Isogai, T.  
 NEDO human cDNA sequencing project  
 Unpublished  
 2 (bases 1 to 2634)  
 TITLE Isogai, T. and Yamamoto, J.  
 JOURNAL Direct Submission  
 COMMENT Submitted (04-JUL-2002) Takao Isogai, FLJ Project (HRI Team): 2-6-7 Kazusa-Kametani, Kisarazu, Chiba 292-0812, Japan (E-mail: genomics@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986) NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan: cDNA full insert sequencing: Research Association for Biotechnology (RAB): cDNA library construction: Helix Research Institute (HRI) (supported by Japan Key Technology Center etc.): 5'- & 3'-end one pass sequencing: RAB, HRI, and Biotechnology Center, National Institute of Technology and Evaluation: clone selection for full insert sequencing: HRI and RAB: annotation: HRI and RAB.  
 FEATURES  
 source Location/Qualifiers  
 1..2634  
 /organism="Homo sapiens"

Query Match 31.5%; Score 450.6; DB 9; Length 2634;  
 Best Local Similarity 64.8%; Pred. No. 1.8e-77;  
 Matches 703; Conservative 0; Mismatches 374; Indels 8; Gaps 2;  
 /db\_xref="taxon:9606"  
 /clone="DNE52007332"  
 /cell\_type="normal dermal fibroblasts (Neonatal Skin) (NHDF2564)"  
 /clone\_1lb="DNE52"  
 /note="cloning vector: PME18FLJ3-primary culture, normal dermal fibroblasts"  
 BASE COUNT 793 a 580 c 586 g 675 t  
 ORIGIN  
 Oy 187 GAGCGCCGCGGAGTAACTTACGCTGGGTTTCAAGCCGCTCTAGGAGTGGCAGCAAC 246  
 Db 365 GACAGCGAGAGATGAGGCGCATCTGTAATTCGAATCTTTTCCAAAGAGAGAGAGACC 424  
 Oy 247 ATGGAACACTTGGAGAACAGCTCATTTGCCCATCTGCTGAGAGATGTTTCGAAGCC 306  
 Db 425 ATGGAATTAATTAGAGAACCACTATCTGTCCTCATCTGTTAGAGATGTTTCGAACCT 484  
 Oy 307 GTGATGATCTTGCCTGCAACACCACTGTGCGGCGCAAGTGTGCCAGACGCTTCCAG 366  
 Db 485 GTGATGATCTTGCCTGCAACACCACTGTGCGGCGCAAGTGTGCCAGACGCTTCCAG 544  
 Oy 367 GCCTTAATCTCTGTGGCAATCCCGGGCTCCACAGCGGTCTTCAAGAGAGCTTTC 426  
 Db 545 GCCTTAATCTCTGTGGCAATCCCGGGCTCCACAGCGGTCTTCAAGAGAGCTTTC 604  
 Oy 427 CGATGCCATCTGTGAGCAGAGGTTGCTCTGAGCAGGATGTTGTATGCTGAG 486  
 Db 605 GCCTGCCATCTGTGAGCAGAGGTTGCTCTGAGCAGGATGTTGTATGCTGAG 664  
 Oy 487 CGGAACCTGTAGTGAAGAACATGATGATCTTACAGCAGAGAGTCTCCCGGCGCTG 546  
 Db 665 AGGAACCTGTGTGGAGAAATATATCATTTGATGATCTTACAGCAGAGAGTCTCCCGGCGCTG 721  
 Oy 547 CAGGCCAAGCTGAGAACAGCACCTTATGTGAGAGAGCAGAGAGAGATCAACATC 606  
 Db 722 --GAAAGAAATCCGACACAGCCCATGTGCGAGAGACATGAAAGAGAGCGCATCAACATC 778  
 Oy 607 TACTGCTGAGCTGCGAGAGTGGCCAGCTGCTCTGCAAGGTTTGGGGCCACAGAG 666  
 Db 779 TACTGCTGAGCTGCGAGAGTGGCCAGCTGCTCTGCAAGGTTTGGGGCCACAGAG 838  
 Oy 667 GACTGTGAGGTGGCCCTCTGCCACCATTTTACAAAGCGAGAGAGTGAAGTGAAGCAT 726  
 Db 839 GACTGTGAGGTGGCCCTCTGCCACCATTTTACAAAGCGAGAGAGTGAAGTGAAGCAT 898  
 Oy 727 GGCATGCGCATGTGCTGGGGGGGGAATGACCGTGTGAGCAGCATGATCAGCATGGAG 786  
 Db 899 GGCATGCGCATGTGCTGGGGGGGGAATGACCGTGTGAGCAGCATGATCAGCATGGAG 958  
 Oy 787 GAGGTGTGCCAGACCATTTGAGACCAACAGCGGAGAGAGAGAGTGAAGTGAAGCAT 846  
 Db 959 GACACCTGCAAAATCTATGAGAGATGTTGAGAAAGAGAGAGAGAGAGTGTGAGAGAG 1018  
 Oy 847 TTGGAACCTGTGGCGGTTTGGAGAGAGCGAAGGGCGAAGTCTTCAAGCATGGCC 906  
 Db 1019 TTTGATTACCTGATGATGATTTTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1078  
 Oy 907 CGGAGACAGAGAGAGAGAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 966  
 Db 1079 CGAACCAGAGAGAGAGAGAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1138  
 Oy 967 TTGAGAGGCTCTCAAAAGCTGTGAGAGTCCGATCCATGATGAGAGAGAGAGAGAGAG 1026  
 Db 1139 TTGAGAGAGCTCTCAAAAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1198  
 Oy 1027 GCTCTACTCTCCAGCAGCAGCAAGAGCTGATCAACAGAGTGGGGCAATGTGAAGTG 1086  
 Db 1199 GCAGTGTCTCAGATATCCCAAAACCTGCTGTAATAAAATTCGGAAGAGATCAAGAGCA 1258







TITLE  
JOURNAL  
REFERENCE  
AUTHORS

TITLE  
JOURNAL

COMMENT

Perera,A., Shim,C., Thomas,E. and Kucherlapati,R.  
High Throughput Mouse Sequencing  
Unpublished  
2 (bases 1 to 209211)  
Han,J., Montgomery,K.T., Grills,G., Chiu,D., Decker,J., Fusina,M.,  
Goltz,J., Halder,A., Hall,L., Ioshikhes,I.P., Lee,E., Long,J.,  
Perera,A., Shim,C., Thomas,E. and Kucherlapati,R.  
Direct Submission  
Submitted (29-NOV-2000) Department of Molecular Genetics, Albert  
Einstein College of Medicine Genome Center, 1300 Morris Park Ave.,  
Bronx, NY 10461, USA  
On Dec 21, 2001 this sequence version replaced g1:15148085.  
-----Genome Center  
Center: Harvard Partners Genome Center  
Center Code: HPGC  
Web site: http://www.hpcg.org/Sequence/mouse.html  
Contact: hpcg@mendel.mgh.harvard.edu  
-----Summary Statistics  
Center project name: ABD  
Sequencing vector: pUC18: L08752  
Chemistry: Dye-terminator Big Dye, 100%  
\*Consensus quality: 204023 at least Q20  
\*Consensus quality: 203326 at least Q30  
\*Consensus quality: 201766 at least Q40  
\*Estimated insert size: agarose-FP - N/A  
\*Estimated insert size: 208811 - sum-of-contigs  
Quality coverage: agarose-FP - N/A  
Quality coverage: 10.2 x in Q20 bases; sum-of-contigs estimation  
-----  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 21 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.  
1 30155: contig of 30155 bp in length  
\* 30156 30175: gap of unknown length  
\* 30176 76428: contig of 46253 bp in length  
\* 76429 76448: gap of unknown length  
\* 76449 99269: contig of 22821 bp in length  
\* 99270 99289: gap of unknown length  
\* 99290 119029: contig of 19740 bp in length  
\* 119030 119049: gap of unknown length  
\* 119050 133795: contig of 14746 bp in length  
\* 133796 133815: gap of unknown length  
\* 133816 147516: contig of 13701 bp in length  
\* 147517 147536: gap of unknown length  
\* 147537 161508: contig of 13972 bp in length  
\* 161509 161528: gap of unknown length  
\* 161529 171555: contig of 10027 bp in length  
\* 171556 171575: gap of unknown length  
\* 171576 182235: contig of 10660 bp in length  
\* 182236 182255: gap of unknown length  
\* 182256 190163: contig of 7908 bp in length  
\* 190164 190183: gap of unknown length  
\* 190184 194456: contig of 4273 bp in length  
\* 194457 194476: gap of unknown length  
\* 194477 199081: contig of 4585 bp in length  
\* 199082 201315: contig of 2234 bp in length  
\* 201316 201335: gap of unknown length  
\* 201336 201781: contig of 446 bp in length  
\* 201782 201801: gap of unknown length  
\* 201802 203287: contig of 1486 bp in length  
\* 203288 203307: gap of unknown length  
\* 203308 204141: contig of 834 bp in length  
\* 204142 204161: gap of unknown length  
\* 204162 204246: contig of 85 bp in length  
\* 204247 204266: gap of unknown length  
\* 204267 205695: contig of 1429 bp in length  
\* 205696 205715: gap of unknown length

# FEATURES

source  
1. 209211  
/organism="Mus musculus"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/chromosome="5"  
/clone="RP23-154K6"  
/sex="male"  
1. 30155  
/note="assembly\_name:Contig39"  
30176. 76428  
/note="assembly\_name:Contig38"  
76449. 99269  
/note="assembly\_name:Contig37"  
99290. 119029  
/note="assembly\_name:Contig36"  
119050. 133795  
/note="assembly\_name:Contig35  
clone\_end:r7  
vector\_side:right"  
133816. 147516  
/note="assembly\_name:Contig34"  
147537. 161508  
/note="assembly\_name:Contig33"  
161529. 171555  
/note="assembly\_name:Contig32  
clone\_end:SP6  
vector\_side:right"  
171576. 182235  
/note="assembly\_name:Contig31"  
182256. 190163  
/note="assembly\_name:Contig30"  
190184. 194456  
/note="assembly\_name:Contig29"  
194477. 199081  
/note="assembly\_name:Contig28"  
199082. 201315  
/note="assembly\_name:Contig27"  
201336. 201781  
/note="assembly\_name:Contig26"  
201802. 203287  
/note="assembly\_name:Contig25"  
203308. 204141  
/note="assembly\_name:Contig24"  
204162. 204246  
/note="assembly\_name:Contig23"  
204267. 205695  
/note="assembly\_name:Contig22"  
205716. 206923  
/note="assembly\_name:Contig21"  
206944. 207841  
/note="assembly\_name:Contig20"  
207862. 209211  
/note="assembly\_name:Contig19"

BASE COUNT 55539 a 49732 c 50169 g 53288 t 483 others  
ORIGIN

Query Match 25.6%; Score 367; DB 2; Length 209211;  
Best Local Similarity 100.0%; Pred.No. 2.3e-61;  
Matches 367; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAGGAGTGTAGACAGAGTCTCGGAATATGACAGGGGTGAGAGAGCTTTAGGGGAGG 60  
|||||  
Db 190769 AAGGAGTGTAGACAGAGTCTCGGAATATGACAGGGGTGAGAGAGCTTTAGGGGAGG 190828  
|||||  
OY 61 GACAGAGCTCTCCAGAGGAGGAGCAATAGCCGGGATCCCAAGATCCAGCAGCTTAAC 120  
|||||  
Db 190829 GACAGAGCTCTCCAGAGGAGGAGCAATAGCCGGGATCCCAAGATCCAGCAGCTTAAC 190888

```
OY 121 TGACCGAGGAAGGCTGCACAGGCGAGAGGGAAGGCCAACGACAGGGCCACAGCGAGCGAG 180
|||
Db 190889 TGACCGAGGAAGGCTGCACAGGCGAGGGAAGGCCAACGACAGGGCCACAGCGAGCGAG 190948
|||
OY 181 GCTCCAGAGCGCGCGGAGTGAACCTTCACGCTGGGTTTCAAGCCGCTGCTAGGGGATGCG 240
|||
Db 190949 GCTCCAGAGCGCGCGGAGTGAACCTTCACGCTGGGTTTCAAGCCGCTGCTAGGGGATGCG 191008
|||
OY 241 CACAACATGGACAACTTGGAGAGCAGCTCATTTGCCCACTCTGCTGAGATGTTCTCC 300
|||
Db 191009 CACAACATGGACAACTTGGAGAGCAGCTCATTTGCCCACTCTGCTGAGATGTTCTCC 191068
|||
OY 301 AAGCCGCTGATCTTGCCTGCCCAACACAACTGTGCGCAAGTGTGCCAAGAGCTC 360
|||
Db 191069 AAGCCGCTGATCTTGCCTGCCCAACACAACTGTGCGCAAGTGTGCCAAGAGCTC 191128
|||
OY 361 TTCCAGG 367
|||
Db 191129 TTCCAGG 191135
```

Search completed: December 3, 2002, 10:09:28  
Job time : 3356 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd

OM nucleic - nucleic search, using sw model

Run on: December 3, 2002, 02:19:40 ; Search time 265 Seconds  
(without alignments)  
12160.791 Million cell updates/sec

Title:	US-09-908-988B-1
Perfect score:	1431
Sequence:	1 aagagctgtacacagagtgt.....ataaagactcaagtgtccc 1431

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

```
Minimum DB seq length: 0
Maximum DB seq length: 20000000000
```

Post-processing:	Minimum Match	0%
	Maximum Match	100%
	Listing first	45 summaries

Database : N Geneseq 101.002.\*

1: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT: \*

2: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT: \*

3: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT: \*

4: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT: \*

5: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT: \*

6: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT: \*

7: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT: \*

8: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT: \*

9: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT: \*

10: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT: \*

11: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT: \*

12: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT: \*

13: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT: \*

14: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT: \*

15: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT: \*

16: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT: \*

17: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT: \*

18: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT: \*

19: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT: \*

20: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT: \*

21: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT: \*

22: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT: \*

23: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT: \*

24: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1431	100.0	1431	24	ABA9061	Murine muscle ring
2	924.2	64.6	1913	22	AAH6563	Human protein HP03
3	924.2	64.6	1990	22	AAH78026	Nucleotide sequence
4	861.2	60.2	1349	21	AAAT7433	Human nucleic acid
5	733.4	51.3	2040	24	ABN85313	Human cytoskeleton
6	626.4	43.8	1500	22	AAF27653	DNA encoding human
7	485.2	33.9	2110	22	AAAS42490	Human cDNA encoding
8	484.6	33.9	1597	24	ABA9063	Murine muscle ring
9	482	33.7	1764	22	ABA83058	Human transcriptio

10	480.8	33.6	1231	22	AA5250037	Human bone marrow
11	478.8	33.5	1781	22	AA5250342	Human CDNA encodin
12	450.8	31.5	2590	24	ABA959662	Human cdna encodin
13	439.6	30.7	1762	22	AA5258535	Human cdna encodin
14	419	29.3	1183	22	AAH901117	Human bone marrow
15	390.6	27.3	1796	22	AAH892942	Human bone marrow
16	267.4	18.7	531	21	AACT52825	Human OREFX OREF40
17	267.4	18.7	531	24	ABN781104	Human interleukin
18	229.2	16.0	587	22	AA5253134	Human CDNA encodin
19	162	11.3	446	22	AA5256303	Human CDNA OREF1315
20	161.6	11.3	650	21	AACT51760	Human OREFX OREF1524
21	114.8	8.0	391	21	AACT519695	Human OREFX OREF1524
22	72.4	5.1	494	22	AAH808085	Mutine 7-transmemb
23	70	4.9	3262	21	AAH838303	Human regulation f
24	66.6	4.7	799	19	AAV55831	Nucleotide sequenc
25	66.6	4.7	1926	21	AAAS50284	Epstein Barr virus
26	66.6	4.7	1926	22	AAH22902	EBV tethering prot
27	66.6	4.7	2580	21	AAAT54544	Nucleotide sequenc
28	66.6	4.7	2580	24	AAAT64275	Epstein-Barr Virus
29	66.6	4.7	5452	20	AAAX09933	Anti-sense strand
30	66.6	4.7	8705	20	AA233778	Vector shuttle DN
31	66.6	4.7	9600	19	AAV71693	Vector plasmid PCM
32	66.6	4.7	10380	20	AA22248	Nucleotide sequenc
33	66.6	4.7	10596	14	AAO51731	Plasmid pcISBON f
34	66.6	4.7	10596	17	AAAT40338	Plasmid pcISBON f
35	66.6	4.7	10596	20	AAAT15650	Nucleotide sequenc
36	66.6	4.7	16080	21	AAAS59553	DNA clone pCER Cl.
37	65	4.5	65	24	ABN52379	Mouse spliced tran
38	63.8	4.5	795	19	AAV55830	FLCA insert stabl
39	63.6	4.4	825	22	AAH03886	Human CDNA clone (
40	63.6	4.4	1738	22	AAH15376	Human CDNA sequenc
41	63	4.4	1925	20	AAAX0924	Epstein Barr Virus
42	58	4.1	2004	18	AAAT85356	Nephila clavipes s
43	56	3.5	1588	22	AAH26304	Spider silk protei
44	54.2	3.8	3489	21	AAH02090	Kaposi's sarcoma-a
45	54.2	3.8	3489	22	AAH2901	Nucleotide sequenc

## ALIGNMENTS

RESULT 1	
ABA99061	
ID	ABA99061 standard; DNA; 1431 BP.
XX	
AC	ABA99061;
XX	
DT	15-JUL-2002 (first entry)
XX	
DE	Murine muscle ring finger protein 1 (MURF-1) coding sequence.
XX	
KW	Muscle ring finger; MURF-1; mouse; cardiant; microtubule;
KW	intermediate filament; striated muscle; cardiac hypertrophy;
KW	heart disease; gene; ds.
XX	
OS	Mus musculus.
XX	
FH	Key
FT	Location/Qualifiers
FT	CDS
FT	199..1299
FT	/*tag= a
FT	/product= "MURF-1"
XX	
FN	W0200206318-A2.
XX	
PD	24-JAN-2002.
XX	
PE	18-JUL-2001; 2001WO-US22896.
XX	
PR	18-JUL-2000; 2000US-219020P.
XX	
PA	(TEXA ) UNIV TEXAS SYSTEM.
XX	
PI	Olson EN, Spencer JA;

XX WPI: 2002-241506/29.  
DR P-FSDB: ABB08275.  
XX  
XX Novel muscle ring finger protein useful for drug screening, and for  
PT diagnosing and treating diseases, particularly cardiomyopathies -  
PT  
XX  
XX Claim 4, Page 123-125, 134pp: English.  
XX

line sequence encodes murine muscle ring finger protein 1 (MURF-1). The invention relates to a purified muscle ring finger (MURF) protein, selected from MURF-1, MURF-2 and MURF-3. The MURF proteins of the invention are involved in microtubule and intermediate filament stabilisation of striated muscle cells and have variant activity. The MURF proteins are useful for screening a candidate substance for MURF protein-binding activity, in a cell, cell-free system or *in vivo*, and its effect on interaction of MURF with microtubules, homodimerisation of MURF, MURF-1, MURF-2 or MURF-3 stabilisation of microtubules, interaction of MURF with intermediate filaments, e.g. desmin, vimentin and cytokeratin, and heterodimerisation of MURF. The screened compounds are useful for treating and preventing cardiac hypertrophy and heart diseases. MURF proteins are useful as antigens to immunise animals for the production of antibodies.

Sequence 1431 BP; 338 A; 384 C; 473 G; 236 T; 0 other;

Query Match	100.0%	Score 1431	DB 24	Length 1431
Best Local Similarity	100.0%	Pred. No. 4,5e-310		
Matches 1431	0	Mismatches	0	Indels 0
		Conservative		Gaps 0

QY	1	AAGGAGGTGACAGCAAGTGTCTGGCAATTACACAGGGGTGAGAGAGCTGTTAGGGGAAG	60
QY	1	AAGGAGGTGACAGCAAGTGTCTGGCAATTACACAGGGGTGAGAGAGCTGTTAGGGGAAG	60
QY	61	GACAGCACTTTCCAAAGAGGACCAATATGCGGGATCCCAAGATCCAGTCAGACTAAAC	120
QY	61	GACAGCACTTTCCAAAGAGGACCAATATGCGGGATCCCAAGATCCAGTCAGACTAAAC	120
QY	121	TGACCGAGGAAGGGTGACACAGCGAGGAGCAAGGCCAAGACACAGGGCCACAGGAGGAC	180
QY	121	TGACCGAGGAAGGGTGACACAGCGAGGAGCAAGGCCAAGACACAGGGCCACAGGAGGAC	180
QY	181	GCATCCAGAGCGCCGCGGGATGAACCTTCAGCGTGGGTTCAAGCCGCTGCTAAGGGATGCG	240
QY	181	GCATCCAGAGCGCCGCGGGATGAACCTTCAGCGTGGGTTCAAGCCGCTGCTAAGGGATGCG	240
QY	241	CACAACTATGACCAACTTGGAGAGCAGCACTATTGGCCCCATCGCCGGAGATGTTCTCC	300
QY	241	CACAACTATGACCAACTTGGAGAGCAGCACTATTGGCCCCATCGCCGGAGATGTTCTCC	300
QY	301	AAGCCGCTGGTATCTTTGGCCCTCCCAACACACCTGTGCCGCAAGTGTGCCAAGCAGCTC	360
QY	301	AAGCCGCTGGTATCTTTGGCCCTCCCAACACACCTGTGCCGCAAGTGTGCCAAGCAGCTC	360
QY	361	TTCCAGGCGCTTAATCTCTGTGGCAATCCGGGGCTCCACAACGCTGTCTTCAGAGGA	420
QY	361	TTCCAGGCGCTTAATCTCTGTGGCAATCCGGGGCTCCACAACGCTGTCTTCAGAGGA	420
QY	421	CGTTTCGATGGCCATCTGTAGGCAAGGATTTGTCTGACACAGCATGTGTCTATGCG	480
QY	421	CGTTTCGATGGCCATCTGTAGGCAAGGATTTGTCTGACACAGCATGTGTCTATGCG	480
QY	481	CTGACAGCGAACCCTGCTAAGGAGAACTATTGACATTTACAAAGCAGAGTCTCCCG	540
QY	481	CTGACAGCGAACCCTGCTAAGGAGAACTATTGACATTTACAAAGCAGAGTCTCCCG	540
QY	541	CCACTTCAGAGCCCAAGGCTTAAGACACCTCATGTGTAGAGACACAGAGACAGCAAGATC	600
QY	541	CCACTTCAGAGCCCAAGGCTTAAGACACCTCATGTGTAGAGACACAGAGACAGCAAGATC	600
QY	601	AACATCTACTGCTGAGCTGCGAGGTGCCCACTGCTCTCTGTGCAAGGTTTTCGGCGC	660
QY	601	AACATCTACTGCTGAGCTGCGAGGTGCCCACTGCTCTCTCTGTGCAAGGTTTTCGGCGC	660

QY	661	CCAAAGACCTGTGAGAGTGGCCCCCTTCGCCACACTTTACAAACGCCAGAGAGTGAGCTG	720
Db	661	CACAAGACCTGTGAGTGGCCCCCTTCGCCACACTTTACAAACGCCAGAGAGTGAGCTG	720
QY	721	AGCAGTGGCATTCGCGATGCTGTGTGGCGGGCCATGATCCGCTGTCAGAGCATGTATCAACCCG	780
Db	721	AGCGATGGCATTCGCGATGCTGTGTGGCGGGCCATGATCCGCTGTGTCAGAGCATGTATCAACCCG	780
QY	781	ATGAGAGAGGTTGTCCACAGACCATTGAGGACAACAAGCCGACAGACAGAACTCTTTAAAC	840
Db	781	ATGAGAGAGGTTGTCCACAGACCATTGAGGACAACAAGCCGACAGACAGAACTCTTTAAAC	840
QY	841	CAGAGTTTCGAGACCCGTGTGGCGGTTTTTGGAGAGAGCCAAAGGCGAACTGCTTAAAGCA	900
Db	841	CAGAGTTTCGAGACCCGTGTGGCGGTTTTTGGAGAGAGCCAAAGGCGAACTGCTTAAAGCA	900
QY	901	CTGGCCCGGGAGAGAGAGAAATTTCACAGCGCGCGCGGGGCTTCATCCGCAGTAAAGCA	960
Db	901	CTGGCCCGGGAGAGAGAGAAATTTCACAGCGCGCGCGGGGCTTCATCCGCAGTAAAGCA	960
QY	961	GACCACTTGGAGGGCTCTTAAAGCTGTGAGTCCCGCATCAATCCATGGAGAGGCG	1020
Db	961	GACCACTTGGAGGGCTCTCTAAAGCTGTGAGTCCCGCATCAATCCATGGAGAGGCG	1020
QY	1021	CACATGGCTCTTACCTCCACAGCGCAAGAGAGTGAATCAACAAAGTCTGGGGCAATTGTGC	1080
Db	1021	CAGATGGCTCTTACCTCCACAGCGCAAGAGAGTGAATCAACAAAGTCTGGGGCAATTGTGC	1080
QY	1081	AAGGTGGAGCTGTGACAGAGCGCGGAGCCAGAGCTATATAGAGCATGGAGCAATTCCTCTGTG	1140
Db	1081	AAGGTGGAGCTGTGACAGAGCGCGCGGAGCCAGAGCTATATAGAGCATGGAGCAATTCCTCTGTG	1140
QY	1141	AGCGTGGAGCAGCTGGCCGAATGTTTGGCAACCATGCACTTCCAGCCGGGCGCGCTGGG	1200
Db	1141	AGCGTGGAGCAGCTGGCCGAATGTTTGGCAACCATGCACTTCCAGCCGGGCGCGCTGGG	1200
QY	1201	GATGAAGAGATGACACATAGCTTTGATGGGGAGAGGGCAATGCGGGCTGTGAGAG	1260
Db	1201	GATGAAGAGATGACACATAGCTTTGATGGGGAGAGGGCAATGCGGGCTGTGAGAGAG	1260
QY	1261	GAGCGGCTTGAGCTGCAAAAGGCTCAGGCTCAGCTGACCGCGACTGTGATTCAGAGCGC	1320
Db	1261	GAGCGGCTTGAGCTGCAAAAGGCTCAGGCTCAGCTGACCGCGACTGTGATTCAGAGAGCGC	1320
QY	1321	ACACCCGAAGCGGAGCCAAAGGATGCTGAGATCTGCGCAGAGACCAACCGCGCACCAA	1380
Db	1321	ACACCCGAAGCGGAGCCAAAGGATGCTGAGATCTGCGCAGAGACCAACCGCGCACCAA	1380
QY	1381	GCTGGGCTTCCCGCCCCCGGGAAGGTTCTCAATAAAGAGCTCAAGTGTCCC	1431
Db	1381	GCTGGGCTTCCCGCCCCCGGGAAGGTTCTCAATAAAGAGCTCAAGTGTCCC	1431

RESULT 2	AAH68563
ID	AAH68563 standard; cDNA; 1913 BP.
XX	
AC	AAH68563;
XX	
DT	13-SEP-2001 (first entry)
XX	
DE	Human protein HP03115 coding sequence
XX	
KX	Human; gene therapy; tumour; ss.
XX	
OS	Homo sapiens.
XX	
PN	WO200142302-A1.
XX	
PD	14-JUN-2001.
XX	
FE	06-DEC-2000; 2000MO-JP08631.



```

FT      /*tag= a
FT      /product= "cardiomyopathy associated protein (CAP)"
XX      MO200162767-AI.
XX      30-AUG-2001.
XX      26-FEB-2001; 2001WO-US05888.
XX      24-FEB-2000; 2000US-0184825.
XX      (GENE-) GENE LOGIC INC.
XX      Bednarik D, Greene J, White M;
XX      WPI: 2001-570621/64.
XX      P-PSDB: AAG63832.
XX      Nucleic acid encoding a cardiomyopathy associated protein that is
XX      differentially expressed in human left ventricle assist device (LVAD)
XX      myocardial biopsy samples, useful for diagnosing, preventing or
XX      treating cardiomyopathy -
XX      Claim 2; Page 49-51; 55pp; English.
XX      The present sequence encodes a human cardiomyopathy associated
XX      protein (CAP) that is differentially expressed in human left ventricle
XX      assist device (LVAD) myocardial biopsy samples. The predicted
XX      isoelectric point of CAP is approximately 12.2. The CAP polynucleotide
XX      and polypeptide are useful for diagnosing, preventing or treating
XX      cardiomyopathy. The treatment diminishes the occurrence of at least
XX      one of the following symptoms associated with cardiomyopathy: reduced
XX      ejection fraction, increased left ventricular diastolic dimension,
XX      decreased ventricular wall thickness, increased atrial size, valvular
XX      regurgitation, exertional intolerance or ventricular tachyrrhythmia.
XX      Modulators of CAP are also useful for preventing or treating
XX      cardiomyopathy.
XX      Sequence 1990 BP; 524 A; 535 C; 613 G; 318 T; 0 other:
XX
Query Match      64.6%; Score 924.2; DB 22; Length 1990;
Best Local Similarity 84.9%; Pred. No. 7.4e-197;
Matches 1117; Conservative 0; Mismatches 163; Indels 35; Gaps 6;

```

```

Db      621 CAGGCGGTGACCTCCAGAGCGTGCAGACAGCTCATGTGCGAGACATGAAGAGAGAA 680
QY      597 GATCAACATTTACTGCTGAGCTGCGAGGTGCCCACTGCTCTCTGCAAGTTTTCGG 656
Db      681 GATCAATATTTACTGCTGAGCTGTGAGTGGCCACCTGCTCTCTCTGCAAGTTTTCGG 740
QY      657 CGCCCAAGAGAGCTGTGAGTGGCCCTGTGCCACCATTTTCAAAAGCCCAAGAGTGA 716
Db      741 TGCCCAAGAGAGCTGTGAGTGGCCCACTGCGACCATTTTCAAAAGCCCAAGAGTGA 800
QY      717 GCTAGCGATGAGCTGCGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 776
Db      801 GCTAGCGATGAGCTGCGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 860
QY      777 CCAGATGAGAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 836
Db      861 ACAGATGAGAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 920
QY      837 AACCCAGAGTGTGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 896
Db      921 AACCCAGAGTGTGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 980
QY      897 AGCACTGCGCCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 956
Db      981 GCGCTGCGCCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1040
QY      957 CGAGAGACACTTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1016
Db      1041 TGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1100
QY      1017 GCGCAGATGAGCTGCTTACCTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1076
Db      1101 GCGCAGATGAGCTGCTTACCTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1160
QY      1077 GTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1136
Db      1161 GTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1220
QY      1137 TGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1196
Db      1221 GGTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1280
QY      1197 TGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1256
Db      1281 CGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1334
QY      1257 GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1316
Db      1335 GGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1388
QY      1317 GCGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1376
Db      1389 GCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1434
QY      1377 CCAAGCTCGGCTTCCGCGCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1431
Db      1435 CCCAAATCGG---CGCGGCGCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1486

```

```

RESULT 4
AAAT2433
ID      AAAT2433 standard; cDNA; 1349 BP.
XX
XX      AAAT2433;
XX
XX      19-DEC-2000 (first entry)
XX
XX      Human nucleic acid-binding protein NuBP-52 cDNA.
XX
XX      Human nucleic acid-binding protein; NuBP; agonist; antagonist; EST;
XX      expressed sequence tag; drug screening; recombinant expression; antibody;
XX      reproductive disorder; infertility; immunological disorder;
XX
KW

```

KM neurological disorder; cell proliferative disorder; cancer; tumour; ss.  
 XX Homo sapiens.  
 OS  
 XX MO200044900-A2.  
 PN  
 XX 03-AUG-2000.  
 PD  
 XX 28-JAN-2000; 2000MO-US02237.  
 PF  
 XX 29-JAN-1999; 99US-0117904.  
 PR 29-JAN-1999; 99US-0117905.  
 XX  
 PA (INCYT-) INCYTE PHARM INC.  
 XX  
 PI Tang YT, Lal P, Hillman JL, Yue H, Azimzai Y, Lu AMD, Baughn MR;  
 PI Tran B, Shih LL, Au-Young JL;  
 XX  
 DR WPI: 2000-499332/44.  
 DR P-PSDB: AAB21048.  
 XX  
 PS Novel nucleic acid binding proteins, used to identify agonists and  
 PT antagonists of them, for the treatment of reproductive, immunological,  
 PT neurological and cell proliferative disorders including cancer -  
 XX  
 PS Claim 4: Page 178-179; 180pp; English.

Sequences AAA72382-A72436 represent cDNAs encoding novel human nucleic  
 CC acid-binding proteins (NuABPs: AAB20997-B21051). These cDNAs were  
 CC produced by extension from an appropriate EST (expressed sequence  
 CC tag) using primers designed using the EST. The invention also relates  
 CC to expression constructs, host cells and transgenic organisms comprising  
 CC a human NuABP nucleic acid, recombinant production of the human NuABPs,  
 CC and antibodies against the human NuABPs, and also to methods of  
 CC screening modulators of human NuABP activity or expression. The human  
 CC NuABPs, and their agonists and antagonists are used to treat diseases  
 CC associated with overexpression or underexpression of functional NuABPs.  
 CC Human NuABP proteins and nucleotides, and NuABP agonists and antagonists  
 CC can be used to diagnose, treat and prevent reproductive, immunological,  
 CC neurological and cell proliferative disorders. Reproductive disorders  
 CC that may be treated using compositions of the invention include  
 CC infertility, endometriosis, disruptions of the menstrual cycle and  
 CC disruptions of spermatogenesis. Immunological disorders that may be  
 CC treated include AIDS, allergies, and autoimmune disorders such as  
 CC multiple sclerosis, rheumatoid arthritis, diabetes and systemic lupus  
 CC erythematosus. Neurological disorders that may be treated include  
 CC epilepsy, neurodegenerative conditions such as Alzheimer's disease and  
 CC Parkinson's disease, prion diseases such as Creutzfeldt-Jakob disease,  
 CC and mental disorders such as schizophrenia. Cell proliferative disorders  
 CC that may be treated include a wide variety of cancers, and also  
 CC arteriosclerosis, atherosclerosis, cirrhosis and psoriasis.  
 CC  
 XX Sequence 1349 BP; 328 A; 372 C; 428 G; 221 T; 0 other;

Query Match 60.2%; Score 861.2; DB 21; Length 1349;  
 Best Local Similarity 88.2%; Pred. No. 7.9e-183;  
 Matches 961; Conservative 0; Mismatches 123; Indels 6; Gaps 2;

DB 118 AACTGACCGAGGAGGTGACAGCAG-GGAGAGGCGCAACGACAGGCGCCACAGCGAG 176  
 DB 52 ATCTAGCGAGGAGGTCTACAGGCACTGAGTGAAGGCCAGAGGAGGCGCCAGCGCAG 111  
 QY 177 GCAGGCTCCAGAGCGCGCGGATGAACCTTCAGCGTGGTTTCAAGCGCTGCTAGAGGA 236  
 DB 112 GCAGGAC-----ACGAGGGGATGAACCTTCAGCGTGGTTTCAAGCGCTGCTAGAGGA 166  
 QY 237 TGGGCAACAATGACAACTTGGAGAGAGAGCTCAATTTGCCCATCTGCTGAGATGTT 296  
 DB 167 TGCACACAGATGACAACTGAGAGAGAGCTCAATCTGCCCATCTGCTGAGATGTT 226  
 QY 297 CTCGACCGCGGTGGTATCTTGGCGCAACACAACTGCGCGGCGGAGTGGCCACAGA 356  
 DB 227 CTCGACCAAGCTGTGATCTCTGCGCGCAACAACTGTGCGGCAATGTGCCACAGA 286

QY 357 CGTCTTCAGAGCCCTAATCCCTGTGGCAATCCGGGGCTCCACAAAGGATCTTCAGG 416  
 DB 287 CGTCTTCAGAGCCCTGAAATCTCTATGGCAATCCCGGGGCTCCACAAAGGATCTTCAGG 346  
 QY 417 AGAAGTTCCTGATGCCATCTTGTAGGACAGAGGTTGCTCTGAGACAGCATGTGTCTA 476  
 DB 347 AGGCGTTTCCGCTGCCCATCTGTGAGGATGAGGTTGTCTGAGACAGACGGTGTCTA 406  
 QY 477 TGGCTTCAGAGGAGCTGTAGTGGAGAACATCATTTACATTCACAAAGGAGTCTTC 536  
 DB 407 CGGCTTCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 466  
 QY 537 CCGGCGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 596  
 DB 467 CAAGCGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 526  
 QY 597 GATCAACATCTACTGCTGAGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 656  
 DB 527 GATCAATATTTACTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 586  
 QY 657 CGGCGCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 716  
 DB 587 TGCACCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 646  
 QY 717 GCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 776  
 DB 647 GCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 706  
 QY 777 CGAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 836  
 DB 707 ACAGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 766  
 QY 837 AAACCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 896  
 DB 767 AAACCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 826  
 QY 897 AGACATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 956  
 DB 827 GGGGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 886  
 QY 957 CGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1016  
 DB 887 TGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 946  
 QY 1017 GCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1076  
 DB 947 GCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1006  
 QY 1077 GTCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1136  
 DB 1007 GTCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1066  
 QY 1137 TGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1196  
 DB 1067 CGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1126  
 QY 1197 TGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1206  
 DB 1127 CGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1136

RESULT 5  
 ABB85313  
 ID ABB85313 standard; cDNA; 2040 BP.  
 XX  
 AC ABB85313;  
 XX  
 AC 30-SEP-2002 (first entry)  
 XX  
 DE Human cytoskeleton-associated protein, CSAP-4, coding sequence.  
 XX  
 KW Human; cytoskeleton-associated protein; CSAP; CSAP-4;

KW cell proliferative disorder; viral infection; neurological disorder;  
KW transgenic animal; antiatherosclerotic; antipsoriatic; antiinflammatory;  
KW virucide; anticonvulsant; vasotropic; cerebroprotective; nootropic;  
KW neuroprotective; cytoskeletal; gene therapy; gene; ss.  
OS Homo sapiens.  
XX  
XX  
FH Key Location/Qualifiers  
FT CDS 291..1271  
FT /tag= a  
FT /product= "CSAP-4"  
PN WO200253719-A2.  
XX  
XX 11-JUL-2002.  
XX  
XX 04-JAN-2002; 2002WO-0500178.  
XX  
XX 04-JAN-2001; 2001US-260085P.  
PR 13-FEB-2001; 2001US-268554P.  
PR 14-FEB-2001; 2001US-269111P.  
PR 23-FEB-2001; 2001US-271211P.  
XX  
XX (INCY-) INCYTE GENOMICS INC.  
PI Lu DM, Baughn MR, Yao MG, Ding L, Honchell CD, Yue H, Tang YT;  
PI Warren BA, Dugan BM, Xu Y, Malia NK, Griffin JA, Stewart EA;  
PI Gandhi AR, Khan FA, Thangavelu K, Ison CH, Azimzal Y, Hatalia AJA;  
PI Gietzen KJ, Lal PG, Sanjanwala MM, Elliott VS;  
DR MPI: 2002-583611/62.  
DR P-SDB: ABB83475.  
XX  
PT Novel isolated human cytoskeleton-associated protein for diagnosing,  
PT treating or preventing atherosclerosis, psoriasis, leukemia, epilepsy,  
PT ischemic cerebrovascular disease, cerebral neoplasms and Alzheimer's  
PT disease -  
XX  
XX  
PS Claim 5; Page 154-155; 167pp; English.  
XX  
CC The present sequence is the coding sequence for a human  
CC cytoskeleton-associated protein (CSAP). The CSAP and its coding sequence  
CC are useful in the diagnosis, treatment and prevention of a cell  
CC proliferative disorder such as actinic keratosis, atherosclerosis,  
CC psoriasis, primary thrombocythemia, leukemias; a viral infection such as  
CC those caused by adenoviruses (acute respiratory disease, pneumonia),  
CC arenaviruses (lymphocytic choriomeningitis); and a neurological disorder  
CC such as epilepsy, ischemic cerebrovascular disease, stroke, cerebral  
CC neoplasms, Alzheimer's disease, Pick's disease, Huntington's disease or  
CC amyotrophic lateral sclerosis. The CSAP coding sequence is also useful  
CC for creating knock out or knock in humanised animals or transgenic  
CC animals to model human diseases.  
XX  
SQ Sequence 2040 BP; 465 A; 584 C; 650 G; 341 T; 0 other;  
Query Match 51.3%; Score 733.4; DB 24; Length 2040;  
Best Local Similarity 78.2%; Pred. No. 3e-154;  
Matches 1028; Conservative 0; Mismatches 156; Indels 131; Gaps 7;  
QY 118 AACTGACCGAAGAGGTGCACAGCAG-GGAGAGAGGCCACGACAGCGCAG 176  
DB 214 ATCTAAGCGAGGAGGCTCTACAGCAGTGAAGGCGAGGAGCGCCAGCGCAG 273  
QY 177 GCAGGCTCAAGCGCCCGGATGACACTTCACGCTGGTTTAAAGCGCTGCTAGGGA 236  
DB 274 GCACGACC-----ACCGAGGGGATTAATTCACAGTGGGTTTAAAGCGCTGCTAGGGA 328  
QY 237 TGCCACACAACTGACACTTGGAGAGAGCATTCATTCGCCCATCTGCTGAGATGT 296  
DB 329 TGACACACGATGAGCACTGAGAGAGCATCTCTGCGCATCTGCGTGAAGATGT 388  
QY 297 CTCCAGCGCGTGTGATCTTCCCTGCCACAGCAACTGTGCGGCAAGTGTGCAACGA 356  
||||| ||| ||||||||| ||||||||| ||||||||| ||||||||| ||||||||| |||||||||

DB 389 CTCCAAACCACTGTGATCTTCCTGCTGCCAACAACACTCTGCGGCAATGTGCCAACGA 448  
QY 397 CGTCTTCCAGGCTCTAATCTCTGTGGCAATCCGGGGCTCCACACAGGCTTCAAG 416  
DB 449 GCTCTTCCAGGCTCTAATCTCTGTGGCAATCCGGGGCTCCACACACTGTGTCTTCAAG 508  
QY 417 AGGAGCTTTCCATGCGCATCTTGTGTAGGAGAGGTTGTCTGTGACAGGAGTGTGTCTA 476  
DB 509 AGGCGGTTTCCGCTGCGCATCTGTGAGGAGATGAGGTGTCTGTGACAGAGAGTGTCTA 568  
QY 477 TGGCTTGCAGCGGAACCTCTGAGAGAAATCATTCATCTACATCTTCAAGCAGAGTCTCT 536  
DB 569 CGGCTTGCAGCGGAACCTCTGAGAGAAATCATTCATCTTACATCTTCAAGCAGAGTCTCT 628  
QY 537 CCGGCGCTGACGCGCAAGGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 596  
DB 629 CAGGCGCTGACGCGCAAGGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 688  
QY 597 GATCAACATCTACTGCTGAGCTGCGAGAGTGGCCACCTGCTCTCTGCAAGGTTTTCG 656  
DB 689 GATCAATATTTACTGCTGAGCTGTGAGGTGGCCACCTGCTCTCTGCAAGGTTTTCG 748  
QY 657 CGCCCAAGAGAGACTGTGAGTGGCCCTCTGCGCACATTTACAAAGCCGAAGAGTGA 716  
DB 749 TGCCCAAGAGAGACTGTGAGTGGCCCTCTGCGCACATTTACAAAGCCGAAGAGTGA 801  
QY 717 GGTGAGCATGAGCATGCGCATGCTGTGGGGGGAATGACCGTGTGACGAGTGTATC 776  
DB 802 ----- 801  
QY 777 CCAGATGAGAGAGTGTGCGCAGACATTTAGAGACACAGCGCGAGAGAGAGAGAGAGTGT 836  
DB 802 -----AGCAATAATGCGGAGAGAGAGAGAGAGAGTGT 832  
QY 837 AAACCAAGAGTGTGAGTGT 896  
DB 833 AAACCAAGAGTGTGAGTGT 892  
QY 897 AGCAGTGGCGCGGAGTGT 956  
DB 893 GCGCGTGGCGCGGAGTGT 952  
QY 957 CGGAGACACTTGTGAGTGT 1016  
DB 953 TGGCGACCACTGTGAGTGT 1012  
QY 1017 GCGGAGTGT 1076  
DB 1013 GCGGAGTGT 1072  
QY 1077 GTCGAAGTGTGAGTGT 1136  
DB 1073 GTCGAAGTGTGAGTGT 1132  
QY 1137 TGTGAGCTGTGAGTGT 1196  
DB 1133 CTTAAGAGTGTGAGTGT 1192  
QY 1197 TGGGATGAGTGT 1256  
DB 1193 CGGAGTGT 1246  
QY 1257 GAGTGT 1316  
DB 1247 GGAAGTGT 1300  
QY 1317 GCGACACCGGAGTGT 1376  
DB 1301 GCGCGGCTAGAGTGTGAGTGT 1346  
QY 1377 CCAAGTGTGAGTGT 1431  
DB 1347 CCCAATGTGAGTGT 1398  
||||| ||| ||||||||| ||||||||| ||||||||| ||||||||| ||||||||| |||||||||



ID	AAF27653	standard; DNA; 1500 BP.
XX	AAF27653;	
XX	30-MAR-2001	(first entry)
XX	DNA encoding human transcriptional regulator protein #22.	
XX	Human; transcriptional regulator protein; TXRBG; ds.	
XX	Homo sapiens.	
XX	MO200078954-AZ.	
XX	28-DEC-2000.	
XX	15-JUN-2000; 2000MO-US16766.	
XX	18-JUN-1999; 99US-0140109.	
XX	(INCY-) INCYTE GENOMICS INC.	
XX	Lal P, Yue H, Tang YT, Baughn MR, Azimzai Y, Tran B.	
XX	WPI; 2001-041425/05.	
XX	Isolated polypeptide with a human transcriptional regulator protein	
XX	sequence is useful for the diagnosis, prevention and treatment of	
XX	PT disorders associated with the immune, reproductive and cardiovascular	
XX	systems -	
XX	Claim 5; Page 135-136; 142pp; English.	
XX	The present invention relates to human transcriptional regulator	
XX	protein (TXRBG) sequences. The antagonist and an agonist of the proteins	
XX	of the invention are used to treat disorders associated with decreased	
XX	or increased expression or activity of TXRBG.	
XX	Sequence 1500 BP; 401 A; 378 C; 439 G; 282 T; 0 other;	
XX	Query Match 43.8%; Score 626.4; DB 22; Length 1500;	
XX	Best Local Similarity 87.5%; Pred. No. 2.2e-130;	
XX	Matches 721; Conservative 0; Mismatches 96; Indels 7; Gaps	
XX	118 AACGTACCCGAGGAGGAGTGCACAGCAG-GGAGAGAGCCAAAGCAGAGGACACAGCGAG 176	
XX	214 ATCTAAGCGAGGAAGGTTCTTACAGGCACTGATGAGGCGAGAGCAGAGGCCAGCGCAG 273	
XX	177 GCAGGCTCCAGAGCGCCGCGAGATGAACCTTACGGTGGGTTTCAAGCCGCTGAGGGGA 236	
XX	274 GCACGACG-----ACCGAGGGAGTGAACCTTACAGTGGGTTTCAAGCCGCTGAGGGGA 328	
XX	237 TGGGCAACATGATGACAACTTGGAGAGCAGCTCATTTGCCCATCTGCTGGAGATGTT 296	
XX	329 TGACACAGAGATGACAACTCGGAGAGACAGCTCATCTGCCCATCTGCTGGAGATGTT 368	
XX	297 CTCCAAGCCCGTGTGATCTTGGCCCTGCGCAACACAACTGTGCCGCAAGTGTGCCAAGA 356	
XX	389 CTCCAAAACGATGGTGTGATCTGCGCCCTGCCAACAACAACCTGTCCGCAAAATGTGCCAAGA 448	
XX	357 CGCTTTCAGAGCTCTATCTCTGTGGCAATCCCGGGGCTCCACAAAGCTGTCTTCAAG 416	
XX	449 CGCTTTCAGAGCTCTGATCTCTATGAGCAGTCCGGGGCTCCACCACTGTGCTTCAAG 508	
XX	417 AGACAGCTTTCGATGCCCATCTTGTAGCAGAGGTTGCTCTGAGACAGCATGTGTCTA 476	
XX	509 AGGCGCTTTCGCTGCCCATCTGTGAGAGCATGAGGTTTGTCTCGAGACAGCATGTGTCTA 568	
XX	477 TGCCCTGACAGCGAACCCTGTAGTGGAGACATCATTTGACATCTACAGCAGAGATCTTC 536	

Db	569	CGGCGCTGCAGCGAAACCTGCTAGTGGAGAACATTATTGCACTTTACAGCAGAGTATC	628
QY	537	CCGGCCACTGCACGCCCAAGCGCTGAAACAGCACCTCATGTGTGAGGAGACGAGCGAA	596
Db	629	CAGCGCCGCTGCATCCAAAGCGTGGAGCAGCACCTCATGTGGCAGGAGCATGAAGAAGAA	688
QY	597	GATCAACATCTACTGCTGTGAGCTGCCAGAGTGGCCACCTGCTCTCTGTGCAAGTCTTCGG	656
Db	689	GATCAATTATTTCACGCTCGAGCTGTGAGGTGCCACCTGCTCTCTGTGCAAGTCTTCGG	748
QY	657	CGCCCAACAGAGACTGTGAGTGGCGCCCTCTGCCACCATTTACAAACGCCAGAGAGTGA	716
Db	749	TGCCCAACAGAGACTGTGAGTGGCGCCCTCTGCCACCATTTACAAACGCCAGAGAGTGA	808
QY	717	GCTGAGCGATGGCATTCGGATGCTGTGTGGCGGGCAATGACCGTGTGCAGGAGTATCAC	776
Db	809	GCTCAGCGATGGCATTCGGATGCTGTGTGGCGGGCAATGACCGCGTCAAGCAGTATCAC	868
QY	777	CCAGATGGAGAGAGTGTGTCCAGACCATTTGAGGACAAACGCCGAGACAGCAACTGTT	836
Db	869	ACAGATGGAGAGAGTGTGTCCAGACCATTTGAGGACAAATGACCGGAGGCGAAGTGTGTT	928
QY	837	AAACGAGAGTTCGAGACCTGTGTGGCGGTTTGGAGGAGCGCAAGGCGCAACTGCTTCA	896
Db	929	AAACGAGAGTTCGAGACCTGTGTGGCGGTTTGGAGGAGCGCGCAACGTTGACTGTGCA	988
QY	897	AGCACTGGCCCCGGGAGC-AGGAGAGAAATTGCAAGCGCGTGGCG	939
Db	989	GCGGCTGGCCCCGGGAGCAAGCAGGAGCAAGCTTCAMCGCGATTCGG	1032
RESULT 7			
AA	42490		
ID	AA542490	standard; cDNA: 2110 BP.	
AC	AA542490;		
XX			
DT	18-DEC-2001	(first entry)	
XX			
DE		Human cDNA encoding an mdt protein, clone LG:247384.1:2000MAY19.	
XX			
KW		Human; molecules for disease detection and treatment; mdt; ss;	
KW		Antiartherosclerotic; hepatotropic; antipsoriatic; cytostatic;	
KW		immunosuppressive; antidiabetic; antiaslomatic; neuroprotective;	
KW		osteopthic; antiarthritic; cell proliferative disorder;	
KW		leukemia; breast cancer; autoimmune disorder; cancer; adenocarcinoma;	
KW		acquired immunodeficiency syndrome; Addison's disease;	
KW		diabetes mellitus; asthma; multiple sclerosis; osteoarthritis.	
OS		Homo sapiens.	
XX			
PN	WO200162922-A2.		
XX			
PD	30-AUG-2001.		
XX			
PE	21-FEB-2001; 2001WO-US05896.		
XX			
PR	24-FEB-2000; 2000US-0185213.		
PR	16-MAY-2000; 2000US-0205232.		
PR	17-MAY-2000; 2000US-0205285.		
PR	17-MAY-2000; 2000US-0205286.		
PR	17-MAY-2000; 2000US-0205287.		
PR	17-MAY-2000; 2000US-0205323.		
PR	17-MAY-2000; 2000US-0205324.		
XX			
PA	(INCY-) INCYTE GENOMICS INC.		
XX			
PI	Panzer SR, Spiro PA, Banville SC, Shah P, Chalup MS, Chang SC;		
PI	Chen A, D'Sa SA, Amshy S, Dahl CR, Dam TC, Daniels SE;		
PI	Dufour GE, Flores V, Fong WT, Greenawalt LB, Hillman JL, Jones AL;		
PI	Liu TF, Roseberry AM, Rosen BH, Russo FD, Stockdreher TK, Daffo A.		

PI Wright RJ, Yap PE, Yu JY, Bradley DL, Bratcher SR, Chen W;  
PI Cohen HJ, Hodgson DM, Lincoln SE, Jackson S;  
DR WPI: 2001-570631/64.  
P-PSDB: AMU25438.  
XX  
XX New disease detection and treatment molecule polynucleotides and  
PT polypeptides, useful for diagnosis and treatment of arteriosclerosis,  
PT cirrhosis, psoriasis, cancer, autoimmune disorders, diabetes mellitus,  
PT asthma and multiple sclerosis -  
XX  
PS Claim 1: Page 121: 183pp: English.  
XX  
CC The invention relates to novel human molecules for disease  
CC detection and treatment (mddt proteins) and the polynucleotides encoding  
CC them. The MDDT polynucleotides and polypeptides are useful for diagnostic  
CC and therapeutic purposes e.g. to diagnose and treat cell proliferative  
CC disorders (e.g. arteriosclerosis, cirrhosis and breast cancer) cancers (e.g.  
CC adenocarcinoma, leukaemia and breast cancer) autoimmune disorders  
CC (e.g. acquired immunodeficiency syndrome (AIDS) and Addison's disease)  
CC diabetes mellitus, asthma, multiple sclerosis, osteoarthritis, and many  
CC more diseases given in the specification. The present sequence  
CC encodes an mddt protein of the invention.

Sequence 2110 BP: 562 A: 465 C: 613 G: 470 T: 0 other;

Query Match 33.9%; Score 485.2; DB 22; Length 2110;

Best Local Similarity 66.8%; Pred. No. 8.2e-99; Mismatches 348; Indels 3; Gaps 1;

Matches 707; Conservative 0; Mismatches 348; Indels 3; Gaps 1;

QY 226 CTGCTAGGGGATGCCACACATGACAACTTGGAGAGAGCTCATTTGCCCATCTGC 285  
DB 143 CTGATCCAGAGATGGGAATCCCATGGAGAACTTGGAGAGAGCTCATTTGCCCATCTGC 202  
QY 286 CTGAGATGTTCTCCAAAGCCCGTGGATGATCTTCCCTGCCCAACAAACCTGGCCGCAAG 345  
DB 203 CTGAGATGTTTACCAACGCACTGTCTCTCCCTGCCAGACAACTGTGCGGAG 262  
QY 346 TGTGCCAAGCAGCTCTCCAGGCTCTAATCTGTGTGCAATCCCGGGCTCCACAAAG 405  
DB 263 TGTGCCAATGATATTTTCCAGGCTCTAACCCTGATTTGCCCAAGAGAGATCCACC 322  
QY 406 GTGTCTCAGAGAGAGCTTCCGATGCCCATCTTGTAGCAGCAGAGTGTCTCTGAGACG 465  
DB 323 ATGCAATCAGGGGCGGCTCCGCTCCATCTGTAGCATGAATGTGTTGGATAGA 382  
QY 466 CATGCTGTATGCTCTGACGCTGACGCAAGCTGTAGTGAAGACATCTTGTACATCAAG 525  
DB 383 CATGGGTATATGACTTCAAGAGAACTGTGTGTGAGAAACATCTCAATCTACAAA 442  
QY 526 CAGAGATCTCTCCGCGCAGCTGACGCAAGGCTGAACAGCACCTCATGTGTGAGAGCAC 585  
DB 443 CAGAGATCTCTCCAGCTGCGCCCTGAGAGAGGCACT--CACCCCATGTGCAAGAGCAC 499  
QY 586 GAGAGAGAGAAAGATCAATCTACTGCTGAGCTGAGAGTGGAGTCCCATCTGTCTCTGC 645  
DB 500 GAGATAGAGAAATCAATCTACTGCTGAGCTGAGAGTGGAGTCCCATCTGTCTCTGC 559  
QY 646 AAGTTTTCGCGCCCAACAGAGCTGTGAGTGGCCCTCTCCCACTTATCAACAAGC 705  
DB 560 AAGTTTTCGCGCCCAACAGAGCTGTGAGTGGCCCTCTCCCACTTATCAACAAGC 619  
QY 706 CAGAAAGTGAAGCTGAGCGATGCGATGCGATGCTGTGTGCGGCGCAATGACCGTGTGACG 765  
DB 620 CAAAGACTGAACTGAATTAATCTATCTCCATGTGTGTGCGGCGGGAATGACCGTGTGACG 679  
QY 766 GCAGTATCAACCCAGATGAGAGAGGTGTGCAGACATTTGAGAGAAACAGCCGCAACAG 825  
DB 680 ACATATCTACTGCTGAGAGATTTCCCGTGAAGTACCAAGAGAAACAGTCAACAGGTA 739  
QY 826 AAGCAACTGTAAACAGAGAGTGTGAGACCTGTGCGCGGTTTGGAGAGGAGGAGGCG 885  
DB 740 AAGGAGAGCTGAGCCAGAGAGTGTGAGACGTTGTATGCACTCTGATGAGAGAAAGT 799

QY 886 GAACGTCTTCAAGCAGCTGCGCCCGGAGCAGAGAGAAAGTTGACGCGCTGCGGGCTTC 945  
DB 800 GAGTTGCTGACGCGGATCAGCAGAGAGAGAGAAAGTTGACGCGCTTCAGAGGCCCTC 859  
QY 946 ATCCGCGCAGAGAGAGACCACTTGGAGGCTTCCCAAGCTGTGGAGTCCGCTCATCCAG 1005  
DB 860 ATCCAGAGATACCAAGAGAGAGCTGACAGATTCACAAAGCTGTGTGAAGTCCCATTCAG 919  
QY 1006 TCCATGAGAGAGCCGCGAGATGCTCTACCTCCAGCAGGCAAGAGAGCTGATCAACAG 1065  
DB 920 TCCCTGAGAGAGAGCTGCGGAGAGCCACTTCTTGAATGCAACCACTATCAAAAGC 979  
QY 1066 GTGCGGCGCATGTGAGAGTGTGAGCTGCGAGAGAGCGCGGAGCCAGCTATAGACATG 1125  
DB 980 ATTTGAGAGCTTCCAAAGGCTGCGAGCTGCGGAGAGAGAGAGAGGCTTTGAGAAATAG 1039  
QY 1126 GAGCAATTTCTGTGAGCGTGAAGCAGTGGCCGAAATGTTGCGAACCATGACTTCCAG 1185  
DB 1040 GACTTCTTACTTTGATTTAGAGCATATGACAGAGCGCCCTGAGAGCCATTTGAGTGGG 1099  
QY 1186 CCGGCGCGCGCTGGGATGTAAGAGATGACGACATGCTTTGATGAGAGAGGCAAT 1245  
DB 1100 ACAGATGAGAGAGAGAGAGAAATTCATTTGAAGAGAAAGATCAGAAAGAGAGATCCACA 1159  
QY 1246 GCGGGCTGAGAGAGAGAGCGGCTGAGCTGCGCAGAG 1283  
DB 1160 GAAAGGAGAGAGAGAGAGACCACTTAAGAGAGCTGATG 1197

RESULT 8  
ABA99063  
ID ABA99063 standard; DNA; 1597 BP.

XX ABA99063;

DT 15-JUL-2002 (first entry)

DE Murine muscle ring finger protein 3 (MURF-3) coding sequence.

XX Muscle ring finger; MURF-3; mouse; cardiant; microtubule;

KW Intermediate filament; striated muscle; cardiac hypertrophy;

KW heart disease; gene; ds.

XX OS

FT Mus musculus.

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

WO200206318-A2.  
24-JAN-2002.  
18-JUL-2001; 2001WO-US22896.  
18-JUL-2000; 2000US-219020P.  
(TEXA ) UNIV TEXAS SYSTEM.  
Olson EN, Spencer JA;  
WPI: 2002-241506/29.  
P-PSDB: ABB08277.  
Novel muscle ring finger protein useful for drug screening, and for  
diagnosing and treating diseases, particularly cardiomyopathies -  
Claim 4: Page 131-133; 134pp; English.  
The sequence encodes murine muscle ring finger protein 3 (MURF-3). The  
invention relates to a purified muscle ring finger (MURF) protein.





	Matches	705; Conservative	0; Mismatches	350; Indels	3; Gaps	17
QY	226	CTCTCTGAGGGATGCGCACAACATGAGCAACCTTGAGAGAGCAGCTCATTTTGGCCCATCTGC				285
Db	141	CTGATCCAGAGATGGGAATCCCATGGAGAACTTGAGAAACACTGTATCTCGCCCTATCTGC				200
QY	286	CTGAGATGTCTCCCAAGCCCGTGAGATCTTGCCCTGCGCAACACAACTGTGGCCGAG				345
Db	201	CTGGAAATGTTTACCAAGCCAGTGGTCATCTTGCCCTGCGCAACAACTGTGGCCGAG				260
QY	346	TGTGCCAACAGCTCTTCCAGGCGCTCTTAATCCTGTGTGCAATCCCGGGCTCCAAAGC				405
Db	261	TGTGCCAAAGACATCTTCCAGGCTCGCAAAATCCCTACTGTGACCGCCGGGCGAGTCAAGT				320
QY	406	GTTGTCTGAGAGAGAGCTTTCGGATGCCCATCTGTGAAGCAGAGAGTGTCTCTGACAG				465
Db	321	TCCATGTCTGGAGGCGGCTTTCCGCTGCCCCACCTGCGCGCAGAGAGTATCATGTAGTGT				380
QY	466	CATGCTGTCTATGCGCTGCGAGCGGAACCTGCTAGTGGAGACATCATTTGACATCTACAAG				525
Db	381	CACGGAGTGTAGCGGCTGCGAGAGAACTGCTGTGGAGAACATCATGACATCTACAA				440
QY	526	CAGGAGTCTCCCGGCGCACTGTGACCCGCAAGGCTGAAACAGCACACTCATGTGTGAGACAC				585
Db	441	CAGGAGTCTCCAGTGGCGGCGCTGCAAGAGGGCAGT--CACCCCATGTGCAAGAGAGCAC				497
QY	586	GAGGAGCAGAGATCAACATCTACTACTGCTGAGACTGCAAGTCCGACCGCTCTCTGCG				645
Db	498	GAAAGATGAGAAATAACATCTACTGTGTCACTGTGAGAGTCCCACTGCTCCATGTGCG				557
QY	646	AAGGTTTTCCGCGCCCAACAAGACTGTGAGTGGCCCTCTGCGCACCATTTTACAACGC				705
Db	558	AAGGTTTTGGGATCCCAAAAGGCTGCGAGGGTGGCCCATTTGCAAGTGTCTCCAGGGA				617
QY	706	CAGAGAGTGAAGCTGAGCGATGCAATCCGAGATGCTGTGGGGGGAATGACCGTGTGAG				765
Db	618	CAAAAGACTGAACTGAATTAATCTGTATCTTCATGCTGTGTGGCGGGAAATGACCGTGTGAG				677
QY	766	GCATGTATCACCAGATGAGAGAGGTTGTGCCAGACCATTTGAGGACMACGCCGCGACAG				825
Db	678	ACCATCATCATCAGCTGAGAGGATTTCCGCTGAGTGAACCAAGAGAACAGTACACAGSTA				737
QY	826	AAGCAACTGTTTAAACCAAGGTTGCGAGACCGTGTCCGGGTTTTGAGAGACGCAAGGCG				885
Db	738	AAGGAAGAGCTGAGCCCAAAATTTGTACACGTTGTATGTCCATCTGTGATGAAGAAAGT				797
QY	886	GAACTGCTTCAAGCACTGGCCCGGGAGCAGAGAGAAAGTTTGCAGCGGTGCGGGGCGCTC				945
Db	798	GAGTGTCTGACGCGGATCACGAGGAGCAGAGGAAAGAAAGCTTAGTTCATGTAAGGCCCTC				857
QY	946	ATCCCCCAGTACGAGAGACCACTTGGAGGGCTCTCTAAAGCTGTGTGAGTCCGCCATTCAG				1005
Db	858	ATCCAGCAGTATCCAGAGCACTGTGACAAAGTCCAAAGCTGTGGTGAAGAACTGCCATCCAG				917
QY	1006	TCCATGGAGAGCCGCAATGAGCTCTCTTACCTCCAGCAGCGGGAAGAGGAGCTATCAACAG				1065
Db	918	TCCCGTGAAGAGCGCTGGGGGAGCCACACTTCTCTTGACTGTGCCAAGCACTATCAAAAGC				977
QY	1066	GTCGGGCAATGTGAGAGTGAAGCTGTGCAAGCAGAGCGCCGGAAGCAGGCTATGAGAGCATG				1125
Db	978	ATTGTGGAAGCTTCCAAAGGCGCTGCCAGCTGTGGGAAGACAGCAGGGCTTTGAGAACATG				1037
QY	1126	GAGCAATCTCTGTGAGGCGTGGAGACAGCTGTGGCCCAAAATGTTGCCAAGCACTGACTTCCAG				1185
Db	1038	GACTTCTTTTACTTTGTGATTTAGACACATACCAACGCGCCCTGGAAGGCACTTGTGCGG				1097
QY	1186	CCGGGCGCCGCTGGGAGTGAAGAGATGACACATGTGCTTTGATGTGAGGAGAGCGGCAAT				1245
Db	1098	ACAGATGAGGAAGGAAGAAATTCATTGSAAGAAAGAATGACGGAAGAGGAAGAGTCCACA				1157
QY	1246	CGGGGCTGTGAGAGAGAGCGGCTGTGACGTGCCAGAGG				1283
Db	1158	GAAAGGAAGGAAGAACACACAGTAAGAGAGCTGTGAGT				1195

ID	AAH90037	standard	cdna	1231	BP.
XX	AAH90037				
AC	AAH90037				
DT	01-OCT-2001	(first entry)			
XX					
DE	Human bone marrow cDNA, SEQ ID NO: 281.				
XX					
KW	Human; bone marrow; antiinflammatory; cytostatic; neuroprotective;				
KW	antiviral; antibacterial; antifungal; anti-HIV; haemostatic;				
KW	immunosuppressive; gene therapy; cytokine cell proliferation;				
KM	cell differentiation modulator; immune disorder; infection; cancer;				
XX	human immunodeficiency virus; HIV; autoimmune disorder; haemophilia; ss.				
OS	Homo sapiens.				
XX					
PN	WO200153453-A2.				
XX					
PD	26-JUL-2001.				
XX					
PF	23-DEC-2000; 2000WO-US34960.				
XX					
PR	21-JAN-2000; 2000US-0488725.				
PR	25-APR-2000; 2000US-0552317.				
PR	09-JUL-2000; 2000US-0598042.				
PR	19-JUL-2000; 2000US-0620312.				
PR	03-AUG-2000; 2000US-0653450.				
PR	14-SEP-2000; 2000US-0662191.				
PR	19-OCT-2000; 2000US-0693036.				
PR	30-NOV-2000; 2000US-0250583.				
XX					
PA	(HYSE-) HYSEQ INC.				
XX					
PI	Ford JE, Boyle BJ, Tang YT, Liu C, Asundi V, Chen R, Ma Y;				
PI	Ren F, Wang J, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;				
PI	Zhao QA, Zhou P, Drmanac RT;				
XX					
DR	WPI; 2001-488707/53.				
DR	P-PSDB; AAM00918.				
XX					
PT	Novel bone-marrow-expressed polynucleotides and polypeptides, useful				
PT	for treating e.g. cancer and immune deficiency disorders -				
XX					
PS	Claim 1; Page 397; 648pp; English.				
XX					
CC	The present sequence is one of 251 novel human polynucleotides				
CC	expressed in the bone marrow. The polynucleotide and the				
CC	polypeptide encoded by it are useful in the treatment of various				
CC	immune deficiencies and disorders. The deficiencies and disorders may				
CC	be genetic, may be caused by a viral (e.g. HIV), bacterial or fungal				
CC	infection, or may result from an autoimmune disorder, a coagulation				
CC	disorder (e.g. haemophilia), inhibition of tumour cell proliferation,				
CC	suppression of an inflammatory response or treatment of a nervous				
CC	system disorder such as Alzheimer's disease. Detection of the presence				
CC	or increased expression of the polynucleotide or the protein it				
CC	encodes is useful for the diagnosis and/or prognosis of one				
CC	or more types of cancer. The polynucleotide and polypeptide can be				
CC	used as nutritional sources or supplements and in the screening of				
CC	chemical compounds as potential drugs.				
XX					
XX					
SO	Sequence 1231 BP; 332 A; 305 C; 361 G; 233 T; 0 other;				
Query Match	33.6%; Score 480.8; DB 22; Length 1231;				
Best Local Similarity	67.2%; Pred. No. 6.9e-98;				
Matches	696; Conservative 0; Mismatches 337; Indels 3; Gaps 1;				
07	226 CTGCTAGGGGATGCCACACAATGGACAACTGGAGAAACAGCTCATTTGCCCATCTGC	285			
db	167 GTTATTCATGAGGATGGGAAATCCCATGGAGAACTGGAGAGACAGCTCATCTGC	226			

QY 286 CTGAGATGTTCTCCAAAGCCCGTGATCTTGCCCTGCCAACAACACTGTGCCGCAAG 345  
DB 227 CTGGAGATGTTTACCAAGCAGAGTGATCTTCCCGTCCAGACAACCTGTGCCGGAAG 286  
QY 346 TGTGCCAAGCAGCTCTCCAGGCTCTAATCTCTGTGSCAATCCGGGGCTCCACAAG 405  
DB 287 TGTGCCAATGACATCTTCCAGGCTGCAAAATCCCTACTGACCCAGGCGGCGAGCTCAGTG 346  
QY 406 GTGTCTGAGAGGAGCTTCCGATGCCATCTTTAGGACAGAGGTGTCCTGGAGAGG 465  
DB 347 TCCATGTCTGGAGGCGGTTCGCTGCCCTCCACCTGCCACAGAGGTGATCATGTGATGT 406  
QY 466 CAGTGTCTATGAGCTGAGCAGCAGCACTGCTAGTGAGAGAACATCATATGATAGAG 525  
DB 407 CAGGAGGTAGGCGCTGCAAGAGAACCTGCTGTGAGAGAACATCATATGATAGAG 466  
QY 526 CAGGAGTCTCCCGGCGCTGACGCCAGGCTGACACGCACTCATGTGTGAGAGAGC 585  
DB 467 CAGGAGTCTCCAGTCCGCGCGCTGAGAGAG---GGCAGTACACCCCATGTGACAGAGC 523  
QY 586 GAGGAGCAGAAAGTCAACATCTACTGCTGAGCTGGAGGAGGCCACCTGCTCTCTGC 645  
DB 524 GAGAGTGAAGAAATCAACATCTACTGCTGACGCTGAGGTGCCACCTGCTCATGTGC 583  
QY 646 AAGGTTTGGCGCCCAACAAGAGCTGTGAGTGAGCCCTGCTGCCACCATTTACAAGGC 705  
DB 584 AAGGTGTTGGATTCACAAGGCGCTGGAGGTGGCCCATTTGACAGAGTCTTCAGAGGA 643  
QY 706 CAGAGAGTGTAGCTGAGCCGATGCTGCTGCTGCTGGCGGCAATGACCTGTGTGAG 765  
DB 644 CAAMAACGTGAATATCTATCTCTCATGCTGTGCTGGCGGGAATGACCTGTGTGAG 703  
QY 766 GCAGTATACCCAGATGAGAGGAGGTGTGCCAGACATTTGAGCAACAGCCGAGAGAG 825  
DB 704 ACCATATCTACTAGCTGAGAGATTCCTCCGTCGAGTGCACAAAGAGAAACAGTACAGTA 763  
QY 826 AAGCAACTGTTAAACCAAGAGTTCGAGACCTGTGCGCGTTTGGAGAGAGCGCAAGGC 885  
DB 764 AAGGAGAGCTGAGCGAGAAATTTGACAGCTTGTATGTCATCTGATGAGAGAAAGT 823  
QY 886 GAACTGCTTCAAGCACTGCGCGGAGAGAGAGAGAAATTTGACAGCGCTGCGGCGCTC 945  
DB 824 GAACTGCTGAGCGGATCACAGAGAGAGAGAAAGCTTACTTATGAGAGCCCTC 883  
QY 946 ATCCGCGAGTACAGAGAGCACTTGGAGGCTCTCAAGAGTGGTGGAGTCCGCATCCAG 1005  
DB 884 ATCCAGCAGTACAGAGAGAGCTGTGACAGTCCACAAAGCTGTGAGAACTGCCATCCAG 943  
QY 1006 TCCATGAGAGAGCGCAGATGCTCTTACTCCAGCAGAGCAAGAGAGCTGATCAACAG 1065  
DB 944 TCCCTGAGAGAGCTGGGAGGAGCCACTTCTTGACTGCCAAGCACTCATCAAAAGC 1003  
QY 1066 GTGGGGGCAATGTCGAAGGTGAGAGTGGAGAGCGCGGAGCCAGCTATGAGAGCATG 1125  
DB 1004 ATGTGTGAACCTTCCAGAGGCTGCGACACTGGGGAAGACAGAGGCGCTTGGAAACATG 1063  
QY 1126 GAGCAATTTCTGTGAGCAGTGGAGCAGCTGCGCAAAATGTTGCGAAGCATGACTTCAG 1185  
DB 1064 GACTTCTTACTTTGATTTAGACACATAGCAGACGCCCTTGAGAGCCATTTGACTTGGG 1123  
QY 1186 CCGGGGCGCGCTGGGATGAAGAGATGAGACATGCTTTGATGGAGAGAGGCAAT 1245  
DB 1124 ACAGATGAGAGAGAGAAATTCATTGAGAGAAAGATCAGAGAGAGAGAGTCCACA 1183  
QY 1246 GCGGGCTGAGAGAGG 1261  
DB 1184 GAGGAGAGAGAGAG 1199

RESULT 11  
AAS25842  
ID AAS25842 standard; cdna; 1781 bp.

XX AAS25842:  
XX 07-NOV-2001 (first entry)  
DE Human cDNA encoding a novel secreted protein. Seq ID 21.  
DE  
DE  
KW Human; immunosuppressive; antiarthritic; ss; antirheumatic;  
KW cytostatic; cardiant; vasotropic; cerebroprotective; nootropic;  
KW neuroprotective; antibacterial; virucide; fungicide; ophthalmological;  
KW vulnary; secreted protein; rheumatoid arthritis;  
KW hyperproliferative disorder; cardiovascular disorder; cardiac arrest;  
KW cerebrovascular disorder; cerebral ischemia; angiogenesis;  
KW nervous system disorder; Alzheimer's disease; infection; ocular disorder;  
KW corneal infection; wound healing; epithelial cell proliferation;  
KW skin ageing; food additive; preservative; antiproliferative.  
XX Homo sapiens.  
XX WO200155322-A2.  
PD 02-AUG-2001.  
PF 17-JAN-2001; 2001WO-US01341.  
PF  
XX 31-JAN-2000; 2000US-0179065.  
PR 04-FEB-2000; 2000US-0180628.  
PR 24-FEB-2000; 2000US-0184664.  
PR 02-MAR-2000; 2000US-0186350.  
PR 16-MAR-2000; 2000US-0189874.  
PR 17-MAR-2000; 2000US-0190076.  
PR 18-APR-2000; 2000US-0198123.  
PR 19-MAY-2000; 2000US-0205157.  
PR 07-JUN-2000; 2000US-0209467.  
PR 28-JUN-2000; 2000US-0214886.  
PR 30-JUN-2000; 2000US-0215135.  
PR 07-JUL-2000; 2000US-0216647.  
PR 07-JUL-2000; 2000US-0216880.  
PR 11-JUL-2000; 2000US-0217487.  
PR 11-JUL-2000; 2000US-0217496.  
PR 14-JUL-2000; 2000US-0218290.  
PR 26-JUL-2000; 2000US-0220963.  
PR 26-JUL-2000; 2000US-0220964.  
PR 14-AUG-2000; 2000US-0224518.  
PR 14-AUG-2000; 2000US-0224519.  
PR 14-AUG-2000; 2000US-0225213.  
PR 14-AUG-2000; 2000US-0225214.  
PR 14-AUG-2000; 2000US-0225266.  
PR 14-AUG-2000; 2000US-0225267.  
PR 14-AUG-2000; 2000US-0225268.  
PR 14-AUG-2000; 2000US-0225270.  
PR 14-AUG-2000; 2000US-0225447.  
PR 14-AUG-2000; 2000US-0225757.  
PR 14-AUG-2000; 2000US-0225758.  
PR 14-AUG-2000; 2000US-0225759.  
PR 18-AUG-2000; 2000US-0226279.  
PR 22-AUG-2000; 2000US-0226681.  
PR 22-AUG-2000; 2000US-0226688.  
PR 23-AUG-2000; 2000US-0227182.  
PR 23-AUG-2000; 2000US-0227009.  
PR 30-AUG-2000; 2000US-0228924.  
PR 01-SEP-2000; 2000US-0229287.  
PR 01-SEP-2000; 2000US-0229343.  
PR 01-SEP-2000; 2000US-0229344.  
PR 01-SEP-2000; 2000US-0229345.  
PR 05-SEP-2000; 2000US-0229509.  
PR 05-SEP-2000; 2000US-0229513.  
PR 06-SEP-2000; 2000US-0230437.  
PR 06-SEP-2000; 2000US-0230438.  
PR 08-SEP-2000; 2000US-0231242.  
PR 08-SEP-2000; 2000US-0231243.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231413.



	QY	466	CATGTTCTCATGTGGCCCGCAGCCGAACCTGCTAGTGGAAACATCAATTGACATCTACACAAG	525
Db	377	CACGCAATTTACGCCCTGCAGAGGAACCTTGCTGTTGGAAACATCAATCAATCTTCACAAA	436	
QY	526	CAGGAGTCTCCCGGCCACTGCACGCCAAGGCTGAACAGCACCTCATGTGTGAGGAGCAC	585	
Db	437	CAGGAGTGTCTCCAGTAGTGGCCCGCTGCAGAAGGGCAGT ---CACCCCATGTGCAAGGAGCAC	493	
QY	646	AAGTTTTTCGGCCCCCAAGAGACTGTGAGAGTGGCCCTTGCCCTGCCAATTCACAAACGC	705	
Db	554	AAGTGTTTGGGATTCACAAAGGCTCTCGAGATGGCCCATTCGACAGTGTCTTCCAGGSA	613	
QY	706	CAGAGAGTAGAGCTGAGACGATGGCATGCGGATGCTGGTGGCGGGCAATGACCCTGTGCAG	765	
Db	614	CAAAAGACTAAMCTGAATACTATCTCATCTGTAAGTGTGAGGTGGCCCACTGCTCCATGTGTC	673	
QY	766	GCAATGATCAACCCAGATGAGAGAGTGTGTCACAGCATTTGAGAGACAACAGCCGACAGAC	825	
Db	674	ACCATCTACCTAGCTGAGGAGATTTCCCGTCGATGACCAAGAGAAAGCTACCAAGTA	733	
QY	826	AAGCAACTGTTTAAACAGAGAGTTCGAGACCTCTGCGCGGTTTTGGAGAGGCGCAAGGAC	885	
Db	734	AAGGAAGAGCTGAGCCAGAGTGTGACAGSTTGTATGCCATCTCGATGATAGAAAGAAAGT	793	
QY	886	GMACTGTTTAACGACTGCGCCCGGAGCAGAGAGAGAGTTCGACGCCGTGCGGGGCTTC	945	
Db	794	GAGTGTCTGACAGGGATTCACGACGAGGACAGGAAAAAACTTAAGCTTCATCGAGGCCCTC	853	
QY	946	ATCCGCAGTAGACGAGACCACTTGGAGGGGCTCCTCAAGCTGGTGGAGTCCGCCATTCAG	1009	
Db	854	ATCCAGCAGTACCAAGAGACGACTGGACAAAGTCCACAAAGCTGGTGAATCCCATTCAG	913	
QY	1006	TCCATGAGGAGGAGCCGAGATGAGCTCTCTACCTCCAGCAGGCAAGAGAGCTGATCAACAAG	1065	
Db	914	TCCCTGAGACGAGCTGGGGGAGGACCACTCTCTTGTACTGTCCAAAGCACTATCAAAGGC	973	
QY	1066	GTCGGGGGCATGTTCGAAGTGGTGAAGCTGGCAGAGAGCGGCCGAGGCTGATAGAGCATG	1122	
Db	974	ATTGTGGAACTTCCAAAGGGCTGCCAGCTGGGGAGAACAGCAGAGCGGCTTTGAGAACATG	1033	
QY	1126	GAGCAATTCCTGTGAGCGTGGAGCACGTGGCCGAATGTTGGAACCATGCACTTCCAG	1185	
Db	1034	GACTTCTTTACTTTGGATTTTAGAGCAATAGCAGACGCGCCTGAGAGCCATTTACTTTGGG	1099	
QY	1186	CCGGGGCCGCTGGGATGTAAGAGATGACAGCATGGCTTTGGATGGGAGAGGAGCAAT	1245	
Db	1094	ACAGATGAGGAAGAGGAGAAATTCATTGAAGAAAGAGATCAGGAAGAGAAAGATGTCACA	1155	
QY	1246	GCGGGCTGAGAGAGAGAGCGGCTGGACGTGCCAAGAGG	1283	
Db	1154	GAGGGAGGAAGAAAGAGACACAGTAAGAGAGCTGATG	1191	
<hr/>				
RESULT 12				
ABA99062				
ID	ABA99062 standard; DNA; 2590 bp.			
AC	ABA99062;			
XX				
XX	15-JUL-2002 (first entry)			
DT				
DE	Murine muscle ring finger protein 2 (MURF-2) coding sequence.			
XX				
KW	Muscle ring finger; MURF-2; mouse; cardiac; cardiact; microtubule;			
KW	intermediate filament; striated muscle; cardiac hypertrophy;			
KW	heart disease; gene; ds.			
XX				
XS	Mus musculus			

Key	Location/Qualifiers
FF	80..11717
FT	/*tag=a
FT	/product= "MURF-2"
PN	WO200206318-A2.
PD	24-JAN-2002.
XX	
XX	18-JUL-2001; 2001WO-US22896.
PF	
XX	18-JUL-2000; 2000US-219020P.
PR	
XX	(TEXA ) UNIV TEXAS SYSTEM.
PA	
PI	Olson EN, Spencer JA;
DR	WPI: 2002-241506/29.
DR	P-PSDB; ABB08276.
XX	
PT	Novel muscle ring finger protein useful for drug screening, and for
PT	diagnosing and treating diseases, particularly cardiomyopathies -
PS	Claim 4; Page 126-129; 134pp; English.
XX	
CC	The sequence encodes murine muscle ring finger protein 2 (MURF-2). The
CC	invention relates to a purified muscle ring finger (MURF) protein,
CC	selected from MURF-1, MURF-2 and MURF-3. The MURF proteins of the
CC	invention are involved in microtubule and intermediate filament
CC	stabilisation of striated muscle cells and have cardiant activity.
CC	The MURF proteins are useful for screening a candidate substance
CC	for MURF protein-binding activity, in a cell, cell-free system or
CC	in vivo, and its effect on interaction of MURF with microtubules,
CC	homodimerisation of MURF, MURF-1, MURF-2 or MURF-3 stabilisation
CC	of microtubules, interaction of MURF with intermediate filaments,
CC	e.g. desmin, vimentin and cytokeratin, and heterodimerisation of MURF.
CC	The screened compounds are useful for treating and preventing
CC	cardiac hypertrophy and heart diseases. MURF proteins are useful as
CC	antigens to immunise animals for the production of antibodies.
XX	
SQ	Sequence 2590 BP; 703 A; 627 C; 625 G; 635 T; 0 other;
Query Match	31.5%; Score 450.8; DB 24; Length 2590;
Best Local Similarity	63.8%; Pred. No. 4.2e-91;
Matches 702; Conservative	0; Mismatches 392; Indels 6; Gaps 1;
QY	187 GAGCGCCGCGGAGTAACCTTCACGSGGTTCACAGCCGCTCTAGGGGATCGCACAAAC 246
Db	68 GACGGCAGGAATAGACACTCTCTGAAATTCACAGCTTTCTCCAAAGACGACAGACC 127
QY	247 ATGCAACAATTGGAGAGCAGCTCATTTTGGCCCATCTGCTCTGGAGATGTTCTCCAAAGCC 306
Db	128 ATGGATTAACCTTGCAAAAACAACTGATCTGTCTCCCATCTGCTAGAGATGTTCCAGAAAGCTT 187
QY	307 GTGGATGATCTGGCCCTGGCAACAACCGTGGCCGACAGTGGCCAGAGAGCTCTTCAG 366
Db	188 GTGGATCTTTCCTTGGCAGCACAACCTGTGTGACGAAGATGTGCCAGTACATCTTCCAG 247
QY	367 GCGCTTAATCTCTGTGGCAATCCCGGGGGCTCCACAAGGATGTTCTTCAGAGAGAGCTTTTC 426
Db	248 GCGCTTAACCGGACTTATCCCAAGAGAGAGGACACCAAGCTGGATCAAGGGGGCGGCTTC 307
QY	427 CGATGCCCATCTTGTAGGACAGAGGTGTCTCTGGACAGGAGCATGGTCTATAGGCTCGAG 486
Db	308 GCGCTTCCCTCTGGAGCATGAGGTGGTGTGTTAGCAACAACAGGGGCTCTATAGCATGGAG 367
QY	487 CGGAACCTGTAGTGAAGAACATCATTTTACATCTCTCAAGCAGAGAGATCTCTCCCGGACATG 546
Db	368 AGGAACCTGTCTGTGAAGAACATTTTGTGATATCTTCAAGCAGAGATCCACCAAGGCCAGAA 427
QY	547 CACGCCAAGCTGAACAGCACCCTCATGTGTGAGAGGACGAGAGCAGGAAGATCAACATG 606







```

Db 619 AGACTGCCAGTGCGCTCCCTCATCATGTGTTCCAGAGACAGAAAGTCTGAGCTCAGTGA 678
Oy 726 TGGCATGCCGAGTCCGAGTGGCGCGGCAATGACCTGTGACAGGAGTGTACACCCAGATGGA 785
Db 679 TGGCATGCCGAGTCCGAGTGGCGAGACAGATGAGTCCAGAGGAGTGTACAGCGCTGGA 738
Oy 786 GAGAGTGTGAGAGACCATTTGAGACAAACCGCCAGACAGACAGACATGTTAAACAGAG 845
Db 739 AGACACCTGCAAAATATTCAGAGAAATGTTGCAAGAAAACAAACAAACAGCTTTTGAGAA 798
Oy 846 GTTCGAGACCTGTGCGCGGTTTGGAGAGCGCAAGGCGCAACCTGTTCAAGCATGCGC 905
Db 799 GTTGTATTAACCTGTATGAGATTTTGGAGAGAGAAATGAATGACCAAGCATCATTTAC 858
Oy 906 CCGGAGAGAGAGAGAAATGTTGACAGCGCTGCGGCGCTCATCCGCGAGTACAGAGACCA 965
Db 859 CCGAAGCCGAGAGGAGAAATGGAACATGTCTGCTCATCAAAAAGTATTTCTGATCA 918
Oy 966 CTTGAGAGGCTCCCTCAAAAGCTGTGAGTCCGCTATCCAGTCCATGAGAGCGCGAGAT 1025
Db 919 TTTGAGAGAACGTCCTCAAAATGTTGATGAGTACAGAAATTTCAATTTATGATGAGCCAGAAAT 978
Oy 1026 GGCCTCTTACTTCCAGACGCAAAAGAGCTGATCAACAGCTGCGGCAATGTCGAAAGT 1085
Db 979 GCGAGTGTCTTTCGAGAAATGCCAAACCTGCTAAAAAAAATCTCAAGAGCATCAAGAGGC 1038
Oy 1086 GGAAGTGCAGAGAGCGGCGAGAGCGAGTATGAGCATGAGAGATTTCTCTGTGAGCGT 1145
Db 1039 ATTTCAGATGAGGAAATATGAAATGAGTATGAGAAATGAGTAAACCTTCCAGAGTCAACCT 1098
Oy 1146 GAGACACCTGCGCGAAATGTTGCGAATCGAATTCAGCTCCAGCGGCGCCGCTGGGAGTGA 1205
Db 1099 CAATGAGAGAGAAAGATTAATACGTGAATTTGACTTTTACAGAGAAATGAAGATGAAGA 1158
Oy 1206 AGAGATATACGATGCGCTTTGGATGGGAGAGAGGCAATGCGGCGCTGGAGAGAGAGCG 1265
Db 1159 AGAAGAAAGAGGC--GGAGAGAGAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGCA 1216
Oy 1266 GCTGGA 1271
Db 1217 GTAGAA 1222

RESULT 14
AAH90117 standard; cDNA; 1183 BP.
XX AAH90117:
XX 01-OCT-2001 (first entry)
XX Human bone marrow cDNA, SEQ ID NO: 474.
XX Human; bone marrow; antiinflammatory; cytosolic; neuroprotective;
XX antiviral; antibacterial; antifungal; anti-HIV; haemostatic;
XX immunosuppressive; gene therapy; cytokine cell proliferation;
XX cell differentiation modulator; immune disorder; infection; cancer;
XX human immunodeficiency virus; HIV; autoimmune disorder; haemophilia; ss.
OS Homo sapiens.
XX
XX WO200153453-A2.
XX
XX 26-JUL-2001.
XX
XX 23-DEC-2000; 2000WO-US34960.
XX
XX 21-JAN-2000; 2000US-0488725.
XX 25-APR-2000; 2000US-0552317.
XX 09-JUL-2000; 2000US-0598042.
XX 19-JUL-2000; 2000US-0620312.
XX 03-AUG-2000; 2000US-0653450.
XX 14-SEP-2000; 2000US-0662191.

```

```

PR 19-OCT-2000; 2000US-0693036.
PR 30-NOV-2000; 2000US-0250583.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Ford JE, Boyle BJ, Tang YT, Liu C, Asundi V, Chen R, Ma Y;
XX Ren F, Wang J, Werhman T, Xu C, Xue AJ, Yang Y, Zhang J;
XX Zhao QA, Zhou P, Drmanac RT;
XX WPI: 2001-488707/53.
XX P-PSDB: AAM00998.
XX
XX Novel bone-marrow-expressed polynucleotides and polypeptides, useful
XX for treating e.g. cancer and immune deficiency disorders -
XX
XX Claim 1; Page 608-609; 648pp; English.
XX
XX The present sequence is one of 251 novel human polynucleotides
XX expressed in the bone marrow. The polynucleotide and the
XX polypeptide encoded by it are useful in the treatment of various
XX immune deficiencies and disorders. The deficiencies and disorders may
XX be genetic, or may be caused by a viral (e.g. HIV), bacterial or fungal
XX infection, or may result from an autoimmune disorder, a coagulation
XX disorder (e.g. haemophilia), inhibition of tumour cell proliferation,
XX suppression of an inflammatory response or treatment of a nervous
XX system disorder such as Alzheimer's disease. Detection of the presence
XX or increased expression of the polynucleotide or the protein it
XX encodes is useful for the diagnosis and/or prognosis of one
XX or more types of cancer. The polynucleotide and polypeptide can be
XX used as nutritional sources or supplements and in the screening of
XX chemical compounds as potential drugs.
XX
XX Sequence 1183 BP; 294 A; 318 C; 327 G; 244 T; 0 other;
XX
XX Query Match 29.3%; Score 419; DB 22; Length 1183;
XX Best Local Similarity 69.1%; Pred. No. 4.4e-84;
XX Matches 588; Conservative 0; Mismatches 260; Indels 3; Gaps 1;
XX
Oy 226 CTGCTAGGAGGATGCGCAACATGAGACACTTGGAGAGAGAGCTGATTTGCCCATCTGC 285
Db 175 CTGATCCAGAGATGGGAATCCCATGAGAGACTTGGAGAGAGAGAGCTATCTGCCATCTGC 234
Oy 286 CTGAGATGTTCTCCAAAGCCCGGTGATCTTCCCTGCGCAACACACACTGTGCCGCAAG 345
Db 235 CTGAGATGTTTACCAAGCCAGATGTCATCTTCCGCTCCAGCAACACTGTGCCGGAAG 294
Oy 346 TGTGCCAAGACGCTTCCAGCGCCTTAATCTCTGTGCGCAATCCCGGGGCTCCCAANG 405
Db 295 TGTGCCAATGACATCTTCCAGCGCTGCAAAATCCCTACTGTGACCAAGCGGAGCTCAGTG 354
Oy 406 GTGTCTTCAGAGAGAGCTTCCGATGCCATCTGTAGACAGAGAGTGTCTCTGACAGAG 465
Db 355 TCCATGCTGTGAGCGCGCTTTCGCTGCCCACTGCGCCCAAGAGTATCATGATGATCT 414
Oy 466 CATGTGTCTATGCGCTTCGACCGGAGAACCTGCTAGTGGAGAAATCATTTGATCAAG 525
Db 415 CAGGAGGTGTACGCGCTCGCAGAGGAACCTGCTGTGGAGAAATCATGATGATCAAAA 474
Oy 526 CAGGAGTCTCCCGGCGCAGTCCAGCGCAAGGCTGAACGACCCATGATGTGAGAGCAG 585
Db 475 CAGGAGTCTCCAGTGGCGCGCTGCAAGAG--GCGACTGACCCCATGTGCAAGAGGAGC 531
Oy 586 GAGGAGAGAGAGATCAACATCTACTGCTGAGTGCAGAGTGCAGGAGTGCAGTCTCTGTC 645
Db 532 GAGATGAGAGAAATCAACATCTACTGCTGACGTGAGAGTGCAGGAGTGCAGTCTCTGTC 591
Oy 646 AAGCTTTGCGCGCCCAAGAGACTGTAGTGGCGCTCTGCGCCACATTTTCAAGAGC 705
Db 592 AAGGTGTTGGAGATCCAAAGGCTGAGAGGTGGCCCATTCAGAGTCTTCCAGAGG 651
Oy 706 CAGAAGGTAGCTGAGAGAGTGCATGCGGAGTGTGGGAGGAGCAATGAGAGTGTGAG 765
Db 652 CAAAAGCTAAGTGAATCACTGATCTCCATGCTGTGGTGGGAGGAGATGACCGGTGCGAG 711

```











to a pathological condition. Antibodies to the proteins can also be used in alleviating symptoms associated with the disorders and in diagnostic immunoassays e.g. radioimmunoassays or enzyme linked immunosorbant assays (ELISA). Disorders which are diagnosed or treated include autoimmune diseases e.g. rheumatoid arthritis, hyperproliferative disorders e.g. neoplasms of the breast or liver, cardiovascular disorders e.g. cardiac arrest, cerebrovascular disorders e.g. cerebral ischaemia, angiogenesis, nervous system disorders e.g. Alzheimer's disease, infections caused by bacteria, viruses and fungi and ocular disorders e.g. corneal infection. The polypeptides can also be used to aid wound healing and epithelial cell proliferation, to prevent skin aging due to sunburn, to maintain organs before transplantation, for supporting cell culture of primary tissues, to regenerate tissues and in chemotaxis. The polypeptides can also be used as a food additive or preservative to increase or decrease storage capabilities, fat content, lipid, protein, carbohydrate, vitamins, minerals, cofactors and other nutritional components. The present sequence encodes a novel secreted protein of the invention.

Query Match 16.0%; Score 229.2; DB 22; Length 587;  
Best Local Similarity 70.9%; Pred. No. 9e-42;  
Matches 333; Conservative 1; Mismatches 129; Indels 7; Gaps 2;

```
OY 187 GAGCGCGCGGGTGAAGTCAAGCGGTGTTGACGCGCTGCTAGGAGATGCGACAC 246
    || || || || || || || || || || || || || || || || || || || || ||
DB 124 GACAGCGAGAGATRAGCCATCTTGATTAACAATCTTTCCAAAGAGCAGACACC 183
    || || || || || || || || || || || || || || || || || || || || ||
OY 247 ATGACAACTTGAGAGAGAGATCATTTGGCCCATCTGCTGAGATGTTCTCCAGCCC 306
    || || || || || || || || || || || || || || || || || || || || ||
DB 184 ATGGATTAAGTGAAGAGAACTCATCTGCTGCTAGAGATGTTACAGAACTT 243
    || || || || || || || || || || || || || || || || || || || || ||
OY 307 GGGGATCTTGGCCCTCCACAGCAACCTGTCGCGAGGTGCGACAGCTTTCAG 366
    || || || || || || || || || || || || || || || || || || || || ||
DB 244 GGGGATCTTGGCCCTCCACAGCAACCTGTCGCGAGGTGCGACAGCTTTCAG 303
    || || || || || || || || || || || || || || || || || || || || ||
OY 367 GCCTCAATCTCTGTGCAATCCGCGGCTCCACAGAGTGTCTTACAGAGAGAGCTTTC 426
    || || || || || || || || || || || || || || || || || || || || ||
DB 304 GCCTCAATCCGATTTTGGCCACAGAGAGATGACCAATGCGATAGGGGCGCATTC 363
    || || || || || || || || || || || || || || || || || || || || ||
OY 427 CGATGCCATCTTGTAGCAGAGTGTCTTCTGAGAGAGATGTCATAGGCTGAG 486
    || || || || || || || || || || || || || || || || || || || || ||
DB 364 CCTGCCCATCTCTGTAGCAGAGTGTCTTCTGAGAGAGATGTCATAGGCTGAG 423
    || || || || || || || || || || || || || || || || || || || || ||
OY 487 CGGAACCTGTAGTGAAGATCATGATCATACAGAGAGAGATGTCATAGGCTGAG 546
    || || || || || || || || || || || || || || || || || || || || ||
DB 424 AGGAACCTGTGTGGAATATCATGATCATACAGAGAGAGATGTCATAGGCTGAG 480
    || || || || || || || || || || || || || || || || || || || || ||
OY 547 CAGCGCAAGCTGAGACACATCTGTGTAGAGAGAGAGAGAGAGATGATCAATC 606
    || || || || || || || || || || || || || || || || || || || || ||
DB 481 ---GAAAGAAATCCGACAGCCCATGTGGAGAGATGAAAGAGAGCGCATCAATC 537
    || || || || || || || || || || || || || || || || || || || || ||
OY 607 TACTGCTGAGTGGAGAGTGGCCATCTCTCTGCAAGGTTTTCG 656
    || || || || || || || || || || || || || || || || || || || || ||
DB 538 TACTGCTGAGTGGAGAGTGGCCATCTCTCTGCAAGGTTTTCG 586
    || || || || || || || || || || || || || || || || || || || || ||
```

RESULT 19  
AAS26303  
ID AAS26303 standard; cDNA; 446 BP.  
XX  
AC AAS26303;  
XX  
DT 07-NOV-2001 (first entry)  
DE Human cDNA encoding a novel secreted protein, Seq ID 482.  
XX  
XX Human; immunosuppressive; antiarthritic; ss; antirheumatic;  
KW cytostatic; cardiant; vasotropic; cerebroprotective; nootropic;  
KW neuroprotective; antibacterial; virucide; fungicide; opthalmological;  
KW vulnerrary; secreted protein; rheumatoid arthritis;  
KW hyperproliferative disorder; cardiovascular disorder; cardiac arrest;  
KW cerebrovascular disorder; cerebral ischaemia; angiogenesis;

KW nervous system disorder; Alzheimer's disease; infection; ocular disorder;  
KW corneal infection; wound healing; epithelial cell proliferation;  
KW skin ageing; food additive; preservative; antiproliferative.  
OS Homo sapiens.  
PN WO200155322-A2.  
PD 02-AUG-2001.  
XX  
PF 17-JAN-2001; 2001WO-US01341.  
XX  
PR 31-JAN-2000; 2000US-0179065.  
PR 04-FEB-2000; 2000US-0180628.  
PR 24-FEB-2000; 2000US-0184664.  
PR 02-MAR-2000; 2000US-0186350.  
PR 16-MAR-2000; 2000US-0189874.  
PR 17-MAR-2000; 2000US-0190076.  
PR 18-APR-2000; 2000US-0198123.  
PR 19-MAY-2000; 2000US-0205515.  
PR 07-JUN-2000; 2000US-0209467.  
PR 28-JUN-2000; 2000US-0214886.  
PR 30-JUN-2000; 2000US-0215135.  
PR 07-JUL-2000; 2000US-0216647.  
PR 07-JUL-2000; 2000US-0216880.  
PR 11-JUL-2000; 2000US-0217487.  
PR 11-JUL-2000; 2000US-0217496.  
PR 14-JUL-2000; 2000US-0218290.  
PR 26-JUL-2000; 2000US-0220963.  
PR 26-JUL-2000; 2000US-0220964.  
PR 14-AUG-2000; 2000US-0224518.  
PR 14-AUG-2000; 2000US-0224519.  
PR 14-AUG-2000; 2000US-0225213.  
PR 14-AUG-2000; 2000US-0225214.  
PR 14-AUG-2000; 2000US-0225256.  
PR 14-AUG-2000; 2000US-0225257.  
PR 14-AUG-2000; 2000US-0225268.  
PR 14-AUG-2000; 2000US-0225270.  
PR 14-AUG-2000; 2000US-0225447.  
PR 14-AUG-2000; 2000US-0225757.  
PR 14-AUG-2000; 2000US-0225758.  
PR 14-AUG-2000; 2000US-0225759.  
PR 14-AUG-2000; 2000US-0225759.  
PR 18-AUG-2000; 2000US-0226279.  
PR 22-AUG-2000; 2000US-0226681.  
PR 22-AUG-2000; 2000US-0226868.  
PR 22-AUG-2000; 2000US-0227182.  
PR 23-AUG-2000; 2000US-0227009.  
PR 30-AUG-2000; 2000US-0228924.  
PR 01-SEP-2000; 2000US-0229287.  
PR 01-SEP-2000; 2000US-0229344.  
PR 01-SEP-2000; 2000US-0229345.  
PR 05-SEP-2000; 2000US-0229509.  
PR 05-SEP-2000; 2000US-0229513.  
PR 06-SEP-2000; 2000US-0230437.  
PR 06-SEP-2000; 2000US-0230438.  
PR 08-SEP-2000; 2000US-0231242.  
PR 08-SEP-2000; 2000US-0231243.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231413.  
PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0232080.  
PR 08-SEP-2000; 2000US-0232081.  
PR 12-SEP-2000; 2000US-0231968.  
PR 14-SEP-2000; 2000US-0232397.  
PR 14-SEP-2000; 2000US-0232398.  
PR 14-SEP-2000; 2000US-0232399.  
PR 14-SEP-2000; 2000US-0232400.  
PR 14-SEP-2000; 2000US-0232401.  
PR 14-SEP-2000; 2000US-0233063.  
PR 14-SEP-2000; 2000US-0233064.  
PR 14-SEP-2000; 2000US-0233065.  
PR 21-SEP-2000; 2000US-0234223.

PR 21-SEP-2000; 2000US-0234274.  
 PR 25-SEP-2000; 2000US-0234997.  
 PR 25-SEP-2000; 2000US-0234998.  
 PR 26-SEP-2000; 2000US-0235484.  
 PR 27-SEP-2000; 2000US-0235834.  
 PR 27-SEP-2000; 2000US-0235836.  
 PR 29-SEP-2000; 2000US-0236327.  
 PR 29-SEP-2000; 2000US-0236367.  
 PR 29-SEP-2000; 2000US-0236368.  
 PR 29-SEP-2000; 2000US-0236369.  
 PR 29-SEP-2000; 2000US-0236370.  
 PR 02-OCT-2000; 2000US-0236802.  
 PR 02-OCT-2000; 2000US-0237037.  
 PR 02-OCT-2000; 2000US-0237038.  
 PR 02-OCT-2000; 2000US-0237039.  
 PR 13-OCT-2000; 2000US-0239935.  
 PR 13-OCT-2000; 2000US-0239937.  
 PR 20-OCT-2000; 2000US-0240960.  
 PR 20-OCT-2000; 2000US-0241221.  
 PR 20-OCT-2000; 2000US-0241785.  
 PR 20-OCT-2000; 2000US-0241786.  
 PR 20-OCT-2000; 2000US-0241787.  
 PR 20-OCT-2000; 2000US-0241809.  
 PR 20-OCT-2000; 2000US-0241809.  
 PR 20-OCT-2000; 2000US-0241826.  
 PR 01-NOV-2000; 2000US-0246517.  
 PR 08-NOV-2000; 2000US-0246474.  
 PR 08-NOV-2000; 2000US-0246475.  
 PR 08-NOV-2000; 2000US-0246476.  
 PR 08-NOV-2000; 2000US-0246477.  
 PR 08-NOV-2000; 2000US-0246478.  
 PR 08-NOV-2000; 2000US-0246523.  
 PR 08-NOV-2000; 2000US-0246524.  
 PR 08-NOV-2000; 2000US-0246525.  
 PR 08-NOV-2000; 2000US-0246526.  
 PR 08-NOV-2000; 2000US-0246527.  
 PR 08-NOV-2000; 2000US-0246528.  
 PR 08-NOV-2000; 2000US-0246532.  
 PR 08-NOV-2000; 2000US-0246609.  
 PR 08-NOV-2000; 2000US-0246610.  
 PR 08-NOV-2000; 2000US-0246611.  
 PR 08-NOV-2000; 2000US-0246613.  
 PR 17-NOV-2000; 2000US-0249207.  
 PR 17-NOV-2000; 2000US-0249208.  
 PR 17-NOV-2000; 2000US-0249209.  
 PR 17-NOV-2000; 2000US-0249210.  
 PR 17-NOV-2000; 2000US-0249211.  
 PR 17-NOV-2000; 2000US-0249212.  
 PR 17-NOV-2000; 2000US-0249213.  
 PR 17-NOV-2000; 2000US-0249215.  
 PR 17-NOV-2000; 2000US-0249216.  
 PR 17-NOV-2000; 2000US-0249217.  
 PR 17-NOV-2000; 2000US-0249218.  
 PR 17-NOV-2000; 2000US-0249245.  
 PR 17-NOV-2000; 2000US-0249264.  
 PR 17-NOV-2000; 2000US-0249265.  
 PR 17-NOV-2000; 2000US-0249297.  
 PR 17-NOV-2000; 2000US-0249299.  
 PR 01-DEC-2000; 2000US-0250160.  
 PR 01-DEC-2000; 2000US-0250391.  
 PR 05-DEC-2000; 2000US-0251030.  
 PR 05-DEC-2000; 2000US-0251988.  
 PR 06-DEC-2000; 2000US-0251479.  
 PR 08-DEC-2000; 2000US-0251856.  
 PR 08-DEC-2000; 2000US-0251868.  
 PR 08-DEC-2000; 2000US-0251869.  
 PR 08-DEC-2000; 2000US-0251889.  
 PR 08-DEC-2000; 2000US-0251990.

PR 11-DEC-2000; 2000US-0254097.  
 PR 05-JAN-2001; 2001US-0259678.  
 XX  
 XX  
 PA (HUMA-) HUMAN GENOME SCI INC.  
 XX  
 PI Rosen CA, Barash SC, Ruben SM;  
 XX  
 XX WPI: 2001-488783/53.  
 DR P-PSDB; AAU16316.  
 XX  
 PT New nucleic acid molecules encoding 461 human secreted proteins for  
 PT diagnosing, preventing, treating or ameliorating medical conditions and  
 XX used as food additives or preservatives -  
 PS Claim 1; SEQ ID NO 482; 980bp; English.

CC The invention relates to isolated nucleic acid molecules and their  
 CC encoded secreted proteins. The nucleic acids and proteins are used to  
 CC prevent, treat or ameliorate a medical condition in e.g. humans, mice,  
 CC rabbits, goats, horses, cats, dogs, chickens or sheep. They  
 CC are also used in diagnosing a pathological condition or susceptibility  
 CC to a pathological condition. Antibodies to the proteins can also  
 CC be used in alleviating symptoms associated with the disorders and in  
 CC diagnostic immunoassays e.g. radioimmunoassays or enzyme linked  
 CC immunosorbant assays (ELISA). Disorders which are diagnosed or treated  
 CC include autoimmune diseases e.g. rheumatoid arthritis,  
 CC hyperproliferative disorders e.g. neoplasms of the breast or liver,  
 CC cardiovascular disorders e.g. cardiac arrest, cerebrovascular disorders  
 CC e.g. cerebral ischemia, angiogenesis, nervous system disorders e.g.  
 CC Alzheimer's disease, infections caused by bacteria, viruses and fungi  
 CC and ocular disorders e.g. corneal infection, and many other  
 CC disorders listed in the specification. The polypeptides can also  
 CC be used to aid wound healing and epithelial cell proliferation, to  
 CC prevent skin aging due to sunburn, to maintain organs before  
 CC transplantation, for supporting cell culture of primary tissues, to  
 CC regenerate tissues and in chemotaxis. The polypeptides can also be used  
 CC as a food additive or preservative to increase or decrease storage  
 CC capabilities, fat content, lipid, protein, carbohydrate, vitamins,  
 CC minerals, cofactors and other nutritional components. The present  
 CC sequence encodes a novel secreted protein of the invention.

Query Match 11.3%; Score 162; DB 22; Length 446;  
 Best Local Similarity 72.5%; Pred. No. 8.5e-27;  
 Matches 222; Conservative 0; Mismatches 82; Indels 2; Gaps 1;

OY 226 CTGCTAGGGGATGGGACACATGAGACACTTGAGAGAGAGCTCATTTGCCCATCTGC 285  
 ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 138 CTGATCCAGGATGGGATGCGCATGAGAACTTGGAGAGAGCTGATCTGCTATCTGC 197  
 OY 286 CTGGAGATGTTTCCAAAGCCCGTGTGATCTTGGCTTGCACCAACACTTGGCCGAG 345  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 198 CTGGAGATGTTTACCAAGCCAGTGTGATCTTGGCTTGCACCAACACTTGGCCGAG 257  
 OY 346 TGTGCCACGACGCTTTCACAGCCGCTTAATCTGTGGCAATCCGGGGCTGCACAAG 405  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 258 TGTGCCAATGACATCTTCCAGGCTGCAATCCCTACTGACCAACCCGGGGAGCTCAGTG 317  
 OY 406 GTGTCTTACGAGGAGAGCTTCCGATGCCCATTTGTGTGGACAGAGGTTGTCCTGGAAGG 465  
 || ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 318 TCCATGTCGTGGAAGCCGTTTCCCTGCCACCTGCGCCGACGAGNAGATCATGGAACGT 377  
 OY 466 CATGTTCTTATGGCCTG--CAGCGGAACCTGCTAGTGAAGAACTGATGATCTTACA 523  
 || ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 378 CACGAGGTGACGGGCTGGCAANAGAACTGCTGGTGGAAACATCATCATCATCTACA 437  
 OY 524 AGCAGG 529  
 || |||  
 Db 438 AACCAAG 443

RESULT 20  
 AAC75760  
 ID AAC75760 standard; cDNA; 650 BP.







GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 3, 2002, 02:23:10 ; Search time 60 seconds  
(without alignments)  
7314.241 Million cell updates/sec

Title: US-09-908-988B-1  
Perfect score: 1431  
Sequence: 1 aagagagctagacagagctgtcc...ataaagactcaagctgtccc 1431

Scoring table:  
IDENTITY\_NUC  
Gapop 10.0, Gapext 1.0

Searched: 441362 seqs, 15338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database:

1: /cg2\_6/ptodata/1/ina/5A.COMB.seq:\*  
2: /cg2\_6/ptodata/1/ina/5B.COMB.seq:\*  
3: /cg2\_6/ptodata/1/ina/5A.COMB.seq:\*  
4: /cg2\_6/ptodata/1/ina/5B.COMB.seq:\*  
5: /cg2\_6/ptodata/1/ina/PTUS.COMB.seq:\*  
6: /cg2\_6/ptodata/1/ina/Backfile1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	485.2	33.9	1757	US-09-484-970B-37	Sequence 37, Appl
2	70	4.9	3262	US-09-561-989-9	Sequence 9, Appl
3	66.6	4.7	1926	US-09-249-585A-2	Sequence 2, Appl
4	66.6	4.7	2580	US-09-050-863-2	Sequence 2, Appl
5	66.6	4.7	2580	US-09-359-081-2	Sequence 2, Appl
6	66.6	4.7	5452	US-09-130-114-1	Sequence 1, Appl
7	66.6	4.7	9600	US-08-910-647-1	Sequence 1, Appl
8	66.6	4.7	9600	US-09-620-925-1	Sequence 1, Appl
9	66.6	4.7	10596	US-07-884-811-15	Sequence 15, Appl
10	66.6	4.7	10596	US-08-087-783A-15	Sequence 15, Appl
11	66.6	4.7	10596	US-08-194-088B-15	Sequence 15, Appl
12	66.6	4.7	10596	US-08-194-087-15	Sequence 15, Appl
13	66.6	4.7	10596	US-08-194-087-15	Sequence 15, Appl
14	66.6	4.7	10596	PCT-US93-04648-15	Sequence 15, Appl
15	66.6	4.5	1926	US-09-249-585A-4	Sequence 4, Appl
16	66.6	4.5	1926	US-09-130-114-2	Sequence 2, Appl
17	59.8	4.2	7218	US-08-232-463-14	Sequence 14, Appl
18	54.2	3.8	3489	US-08-728-323A-1	Sequence 1, Appl
19	54.2	3.8	3489	US-09-298-568-1	Sequence 1, Appl
20	54.2	3.8	32207	US-08-770-379-20	Sequence 20, Appl
21	54.2	3.8	32207	US-08-757-669A-20	Sequence 20, Appl
22	54.2	3.8	32207	US-09-230-371A-20	Sequence 20, Appl
23	53	3.7	2338	US-08-425-069-1	Sequence 1, Appl
24	52.4	3.7	2338	US-08-317-844B-1	Sequence 1, Appl
25	52.4	3.7	2793	US-08-209-747-1	Sequence 1, Appl
26	52.4	3.7	2793	US-08-458-288-1	Sequence 1, Appl
27	50.8	3.5	441529	US-09-103-840A-1	Sequence 1, Appl

28	45.6	3.2	1505	US-07-915-246-1	Sequence 1, Appl
29	45.6	3.2	2214	US-08-864-038A-1	Sequence 1, Appl
30	45.6	3.2	3331	US-08-864-038A-2	Sequence 2, Appl
31	45.6	3.2	3331	US-08-864-038A-4	Sequence 4, Appl
32	45.6	3.2	9551	US-08-056-200-93	Sequence 93, Appl
33	45.6	3.2	9551	US-08-800-644-93	Sequence 93, Appl
34	44.6	3.1	1995	US-08-425-069-3	Sequence 3, Appl
35	44.6	3.1	1995	US-08-317-844B-3	Sequence 3, Appl
36	44.4	3.1	4214	US-09-123-135-1	Sequence 1, Appl
37	44.2	3.1	1767	US-07-903-466-2	Sequence 2, Appl
38	44.2	3.1	1767	PCT-US93-05794-2	Sequence 2, Appl
39	44.2	3.1	3018	US-07-903-466-1	Sequence 1, Appl
40	44.2	3.1	3018	PCT-US93-05794-1	Sequence 1, Appl
41	43.6	3.0	4852	US-07-853-913-3	Sequence 3, Appl
42	43	3.0	2830	US-09-010-928B-1	Sequence 1, Appl
43	43	3.0	3223	US-07-980-528-1	Sequence 1, Appl
44	43	3.0	51255	US-08-781-891-209	Sequence 209, Appl
45	42.8	3.0	765	US-09-603-785-524	Sequence 524, Appl

#### ALIGNMENTS

RESULT 1  
US-09-484-970B-37  
Sequence 37, Application US/09484970B  
Patent No. 6426186  
GENERAL INFORMATION:  
APPLICANT: Jones, Karen A.  
APPLICANT: Volkmutz, Wayne  
APPLICANT: Walker, Michael G.  
TITLE OF INVENTION: BONE REMODELING GENES  
FILE REFERENCE: PB-0014 US  
CURRENT APPLICATION NUMBER: US/09/484,970B  
CURRENT FILING DATE: 2000-01-18  
NUMBER OF SEQ ID NOS: 177  
SOFTWARE: PERL Program  
SEQ ID NO 37  
LENGTH: 1757  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc.feature  
OTHER INFORMATION: Incyte ID No. 6426186 247384.1CBL  
US-09-484-970B-37

Query Match	Similarity	Score	DB 4:	Length	1757:
Best local	66.8%	Pred. No. 6.2e-106;			
Matches 707;	Conservative	0;	Mismatches 348;	Indels 3;	Gaps 1;
QY	226	CTGCTAGGAGATCGGACCAACATGGACACTTGGAGAGAGAGCTATTGGCCCATCTGC	285		
DB	143	CTGATCCAGATATGGATGGAATCCATGGAGACTTGGAGAGAGAGCTATTGGATCTGC	202		
QY	286	CTGGAATGTTCTTCCAGACCCCTGGTATCTTGGCCCTGCAACCAACCTGGCCGAG	345		
DB	203	CTGGAATGTTCTTCCAGACCCCTGGTATCTTGGCCCTGCAACCAACCTGGCCGAG	262		
QY	346	TGTCGAACAGACGCTTCCAGACCCCTGATCTTATCTTGGCCATCCGGGCTCCACAG	405		
DB	263	TGTCGAATATATTTTCCAGACCCCTGATCTTATCTTGGCCATCCGGGCTCCACAG	322		
QY	406	GTCCTTCAGAGAGACGCTTCCAGACCCCTGATCTTATCTTGGCCATCCGGGCTCCACAG	465		
DB	323	ATGCAATCAGAGAGAGACGCTTCCAGACCCCTGATCTTATCTTGGCCATCCGGGCTCCACAG	382		
QY	466	CTGCTATCTATGAGCCCTGCAACCCCTGATCTTATCTTGGCCATCCGGGCTCCACAG	525		
DB	383	CATGGGTATATGAGCCCTGCAACCCCTGATCTTATCTTGGCCATCCGGGCTCCACAG	442		
QY	526	CAGAGTCTTCCGAGCCCTGCAACCCCTGATCTTATCTTGGCCATCCGGGCTCCACAG	585		
DB	443	CAGAGTCTTCCGAGCCCTGCAACCCCTGATCTTATCTTGGCCATCCGGGCTCCACAG	499		

QY	586	GAGGACGAGAAATCAACATCTACGCGCTGAGCTCGAGAGTCCACCGCTCTCTGCG	645
Db	500	GAGATAGAGAAATCAACATCTACTGTCTCAGCTGAGGTGCCACCGCTCCATGTCG	559
QY	646	AAGTTTTGGGCGCCCAAGAGACTGTGAGTGGCCCTCTGCCACATTTACAAACGC	705
Db	560	AAGGTGTGTTGGGATCCCAAGGCCCTGCGAGGTGGGCCCATTTGACAGAGTCTTTCCAGGGA	619
QY	706	CAGAAAGTAGAGCTGAGCGAATGGATCGGATGCGTGGTGGGGCGGCAATATACCGTGTGCGAG	765
Db	620	CAAAAGCTAACTGAATTAAGTGTATCTCCATGCTGGTGGGGGAATGACCGTGTGCGAG	679
QY	766	GCAGTGATCACCCAGATGAGAGAGGTGTGCCAGACCAATTTGAGAGCAACGCGCAGACAG	825
Db	680	ACCATCATCACTAGCTGAGAGATTTCCCGTGGATGACCAAGGAACACAGTATACCACGTA	739
QY	826	AACCAACTGTAAACCAAGAGAGTTGGAACACCTGTGCGCGGTTTTGAGAGAGCGCAAGGCG	885
Db	740	AAGGAAGAGCTGAGACCGCAAGATTTGACACGTTGATGCGATCTGGATGAGAAAGAAAGT	799
QY	886	GAACTGCTTCAAGCAGTACGCGCCGGGAGCAGAGAGAGAGTTGACAGCGGTGGCGGGCCCTC	945
Db	800	GAGTTCTGCGAGCGGATCACGAGAGACAGAGAAAAGCTTAGTTATGAGAGCCCTC	859
QY	946	ATCCGCGCAGTACGGAGACCACTTGTGGAGGCTCTCCAAAGCTGTGTGAGTCCGCCATCCAG	1003
Db	860	ATCCACAGTACCAAGAGGAGCGAGCTGGACAAAGTCCAAAGCTGTGTGAAATGCCATCCAG	919
QY	1006	TCCATGAGAGAGCCGAGATAGTGCCTCTACTCCAGACAGCAAGAGACCTGATCAACAG	1063
Db	920	TCCGTGAGAGAGCTGTGGGGAGGCCACTTCTCTTGACTGCCAAGCACTATCAAAAGC	979
QY	1066	GTCGGGGCAATGTCAAGGTGTGAGCTGTGGACAGACGCGCGGAGCCAGCGCTATGAGAGCATG	1123
Db	980	ATTGTGTGAAGCTTCCAAAGGCGCTCCACGCTGGGGAAAGACAGAGCGGCTTTGAGAACATG	1033
QY	1126	GAGCAATTTCTGTGAGCGGTGAGACAGGTGGCGGAATGTGGCAACCATGAGACTTCCAG	1183
Db	1040	GACTTCTTACTTTGGATTTAGAGCATATGACAGACGCCCTTGAAAGCGCATTTGACTTTGGG	1093
QY	1186	CCGGGCGCCGCTGGGGATGAAGAGATGACAGACATGGCTTTGGATGGGAGAGAGGCGCAT	1243
Db	1100	ACAGATGAGAGAGAGAGAAATTCATTTGAGAGAAAGATCAGGAAGAGAAAGAGTCCACA	1153
QY	1246	GCGGGGCTGAGAGAGAGGCGCTGAGACGTCCCAAGAG	1283
Db	1160	GAAAGGAAAGAAAGAGACCACTAAGAGATGTGATG	1197

```

RESULT 2
US-09-561-989-9
; Sequence 9, Application US/09561989
; Patent No. 6468750
; GENERAL INFORMATION:
; APPLICANT: KOLLER, Klaus-Peter
; TITLE OF INVENTION: No. 6468750el Cell Regulation Factor T7020
; FILE REFERENCE: seq
; CURRENT APPLICATION NUMBER: US/09/561,989
; CURRENT FILING DATE: 2000-04-27
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
; LENGTH: 3262
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-561-989-9

Query Match      4.9%, Score 70; DB 4; Length 3262;
Best Local Similarity 47.1%, Pred. No. 1.4e-07;
Matches 214; Conservative 0; Mismatches 240; Indels 0; Gaps 0;

YX 563 AGCACCTTCATCTGTGAGACGACGACGAGACGAGATCAACTTACTTGCCCTGAAGCTGCG 622

```

Db	329	AGCTCCACCTCGACACAGACCACGAGCTGGAGAAACCAAGCATGTACTGCTGCATATGCA	388
QY	623	AGGTGGCCACCTGCTCTCTCTCTCAGAGTTCCTGGGCCCCAAGAAGCTGTAGTGGCCC	682
Db	389	AGATGGCCCGTGTGTACCAAGTCTTTGGAGAGGGGCAACACTCCAGCCACCAATCTCAAGG	448
QY	683	CTCTGCCACCACTTATACAAAGCCAGAGAGAGAGAGCTGAGGATGGCAATGCCATGCTGG	742
Db	449	CTCTGGGGGCCATGTGGAAACTACATAAGACACAGCTCTCCAGAGCGCTGAACGAGACTGT	508
QY	743	TGGCGGCAATGACCGTGTGCAGGCAAGTATACCCAGATGAGAGAGGTGACCAGACCA	802
Db	509	CAGACACGGGCCAAAGAACGACAGAGGTTTCTGTACAGCTCGCAACATGCTCCAGCAGA	568
QY	803	TTGAGACACACAGCCGCGACAGAAAGCACTTTAAACAGAGGTTTGAGACCTGTGCG	862
Db	569	TCCAGAGACACAGTGTGGAGTTTGAAGCTGTCTGTGTGGCCCAATGTGATGCCCTCATCG	628
QY	863	CGGTTTGGAGAGGCGCAAGGCGCAACTGCTTCAAGCAGTGGCCGGGAGACAGAGAGAGA	922
Db	629	ATGCCCTCAACAGAAAGAAAGGCCAGCTGTGGCCCGCGGTCAACAAAGGACATGAGACA	688
QY	923	AGTTGACAGGCGGTGGGGGGCTCAATCCGCCAGTACGGAGACCACTTGGAGGCTCTCAA	982
Db	689	AGCTAAGATGTTTGAATACGATCTCTCACTACAGTGAATATGCCAGACACAG	748
QY	983	AGCTGTGAGATCGGCATTCAGTCATGGAGGA	1046
Db	749	GTCATATGAGTACTGCTTGGAGGCTGATTTAAGGA	782

```

RESULT 3
US-09-249-585A-2
; Sequence 2, Application US/09249585A
; Patent No. 6417002
; GENERAL INFORMATION:
; APPLICANT: Horlick, Robert
; TITLE OF INVENTION: METHOD FOR MAINTENANCE AND SELECTION OF EPISOMES
; FILE REFERENCE: 0867/OD905
; CURRENT APPLICATION NUMBER: US/09/249,585A
; CURRENT FILING DATE: 1999-02-11
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 2
; LENGTH: 1926
; TYPE: DNA
; ORGANISM: Epstein Barr Virus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1926)
; OTHER INFORMATION: coding strand of EBNA-1 DNA
US-09-249-585A-2

```

Query Match	4.7%	Score 66.6	DB 4	Length 1926
Best Local Similarity	45.5%	Pred. No. 7.7e-07		
Matches 237	Conservative 0	Mismatches 284	Indels 0	Gaps 0

QY	760	GTCCAGGCAATGATACCCAGATGGAGGAGGTGTGCCAGACCATTTAGGAGACAAACGCC	819
Db	290	GGGCGAGGACGAGGAGGGGCGAGGACGAGGAGGCGGACGAGACGAGGAGGGCGAGG	349
QY	820	AGACAGAGCAACATGTTTAAACAGAGGTTCCAGAACCTGTGCGCGTTTGGAGGAGCC	879
Db	350	GGCGAGGAGGGGCGAGGAGGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	409
QY	880	AAGGCGGAATGCTTCAAGCACTGGCCGGGAGCGAGGAGAAAGTTGCACGCGTCGGG	939
Db	410	GAGGGGCGAGGAGGCGAGGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	469
QY	940	GGCCTATCCGCCGATGACGAGACCACTTGGAGGGCTCTTAAAGCTGTGTGAATCCGC	999
Db	470	CAGAGCAGGAGGAGGAGGCGAGGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	529



```

; TYPE: nucleic acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-359-081-2

```

```

Query Match      4.7%; Score 66.6; DB 4; Length 2580;
Best Local Similarity 45.5%; Pred. No. 8.5e-07;
Matches 237; Conservative 0; Mismatches 284; Indels 0; Gaps 0;

```

```

QY 760 GTGACAGCAGTGTATCCACCATGAGAGAGGTGTGCGACCATTTGACCAACAGCCGC 819
    |||||
DB 673 GGGCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 732
    |||||
QY 820 AGACAGAGCAACTGTTAAACAGAGTTCGAGAGACCTGTGCGCGTTTGGAGAGAGCC 879
    |||||
DB 733 GGGCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 792
    |||||
QY 880 AAGGCGCAACTGCTTCAAGCACTGCCCGGAGCAGAGAGAGAGAGAGAGAGAGAGAG 939
    |||||
DB 793 GAGGCGCAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 852
    |||||
QY 940 GGCCTCATCCCGCAATGACAGACCACTTGAAGGCTCTCAAGCTGTGAGATCCGCC 999
    |||||
DB 853 CAGGAGCAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 912
    |||||
QY 1000 ATCCAGTCCATGAGAGAGCCGACAGATGAGCTCTCTACCTCCAGCAGGCAAGAGCTGATC 1059
    |||||
DB 913 CAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 972
    |||||
QY 1060 AACAGGTCGGGGCAATGTGCAAGGTGAGCTGCGAGAGAGCGCCGAGGAGGAGGAG 1119
    |||||
DB 973 GAGCAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1032
    |||||
QY 1120 AGCATGGAGCAATTCCTGTGATGAGCTGTGAGCAGCAGTGGCCGAATTTGGCAACATCGAC 1179
    |||||
DB 1033 GGGCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1092
    |||||
QY 1180 TTCAGAGCCGGGCGCGCTGCGAGATGAAAGCATGACGACATGCTTTGGATGGGAGAG 1239
    |||||
DB 1093 CAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1152
    |||||
QY 1240 GGCAGATCGGGGCTGTGAGAGAGAGCGGCTGTGACGTGCCAGA 1280
    |||||
DB 1153 GGGCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1193
    |||||

```

## RESULT 6

```

US-09-130-114-1/c
; Sequence 1, Application US/09130114
; Patent No. 5976807
; GENERAL INFORMATION:
; APPLICANT: Horlick, Robert A.
; APPLICANT: Dama, Bassam B.
; APPLICANT: Robbins, Alan K.
; TITLE OF INVENTION: Eukaryotic Cells Stably Expressing Genes
; TITLE OF INVENTION: From Multiple Transfected Episomes
; FILE REFERENCE: 0867/1D903US1
; CURRENT APPLICATION NUMBER: US/09/130,114
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 5452
; TYPE: DNA
; ORGANISM: YEBA
US-09-130-114-1

```

```

Query Match      4.7%; Score 66.6; DB 2; Length 5452;
Best Local Similarity 45.5%; Pred. No. 1.1e-06;
Matches 237; Conservative 0; Mismatches 284; Indels 0; Gaps 0;

```

```

QY 760 GTGACAGCAGTGTATCCACCATGAGAGAGGTGTGCGACCATTTGAGAGCAACAGCCGC 819
    |||||
DB 2132 GGGCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2073
    |||||
QY 820 AGACAGAGCAACTGTTAAACAGAGTTCGAGAGACCTGTGCGCGTTTGGAGAGAGCC 879
    |||||
DB 2072 GGGCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2013
    |||||
QY 880 AAGGCGCAACTGCTTCAAGCACTGCCCGGAGCAGAGAGAGAGAGAGAGAGAGAGAG 939
    |||||
DB 2012 GAGGCGCAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1953
    |||||
QY 940 GGCCTCATCCCGCAATGACAGACCACTTGAAGGCTCTCAAGCTGTGAGATCCGCC 999
    |||||
DB 1952 CAGGAGCAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1893
    |||||
QY 1000 ATCCAGTCCATGAGAGAGCCGACAGATGAGCTCTCTACCTCCAGCAGGCAAGAGCTGATC 1059
    |||||
DB 1892 CAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1833
    |||||
QY 1060 AACAGGTCGGGGCAATGTGCAAGGTGAGCTGCGAGAGAGCGCCGAGGAGGAGGAG 1119
    |||||
DB 1832 GAGCAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1773
    |||||
QY 1120 AGCATGGAGCAATTCCTGTGATGAGCTGTGAGCAGCAGTGGCCGAATTTGGCAACATCGAC 1179
    |||||
DB 1772 GGGCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1713
    |||||
QY 1180 TTCAGAGCCGGGCGCGCTGCGAGATGAAAGCATGACACATGCTTTGGATGGGAGAG 1239
    |||||
DB 1712 CAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1653
    |||||
QY 1240 GGCAGATCGGGGCTGTGAGAGAGAGCGGCTGTGACGTGCCAGA 1280
    |||||
DB 1652 GGGCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1612
    |||||

```

## RESULT 7

```

US-08-910-647-1
; Sequence 1, Application US/08910647
; Patent No. 6251433
; GENERAL INFORMATION:
; APPLICANT: Zuckermann et al.
; TITLE OF INVENTION: Compositions and Methods for
; TITLE OF INVENTION: Polynucleotide Delivery
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESS: Chiron Corporation
; STREET: 4560 Horton Street
; CITY: Emeryville
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 94608-2916
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/910,647
; FILING DATE:
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Fujita, Sharon M.
; REGISTRATION NUMBER: 38,459
; REFERENCE/DOCKET NUMBER: 1218, 002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (510) 923-2706
; TELEFAX: (510) 655-3542
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9600 base pairs
; TYPE: nucleic acid

```

STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-08-910-647-1

Query Match 4.7%; Score 66.6; DB 4; Length 9600;  
Best local Similarity 45.5%; Pred. No. 1.3e-06;

Matches 237; Conservative 0; Mismatches 284; Indels 0; Gaps 0;

QY 760 GTGCAGGCAATGATCAACCCAGATGAGAGAGTGTGCCAGACATTGAGGACAAACAGCCGC 819  
D 719 GGGCAGGACAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 778  
QY 820 AGACAGCAACTGTTAAACCCAGAGGTTGAGAGCCCTGTCGCGGTTTGGAGAGAGCGC 879  
D 779 GGGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 838  
QY 880 AAGGCGCAACTGCTTCAAGCACTGCCCCGAGCAGAGAGAGAAATTGCAAGCCGCTGCGG 939  
D 839 GAGGGGCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 898  
QY 940 GGCCTCATCCGCCAGTACGAGACCACTTGGAGGCGTCTCAAACTGCTGAGATCCGCC 999  
D 899 CAGGAGCAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 958  
QY 1000 ATCCAGTCCATGAGAGAGCCGACAGTGGCTCTTACCTCCAGAGGCAAAAGAGCTGATC 1059  
D 959 CAGGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1018  
QY 1060 AACAAAGTCCGGGCGCAATGTCGAAGGTGAGCTGCAAGAGCGCCGAGACCCAGGCTATGAG 1119  
D 1019 GAGCAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1078  
QY 1120 AGCATGAGCAATTCCTGTGATGAGCTGAGACACGTGGCCGAATTTTGGAAACCATCGAC 1179  
D 1079 GGGCAGGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1138  
QY 1180 TTCAGCCGGGCGCGCTGGGATGAAGAGATGACGACATGCTTTGGATGGGAGAGAG 1239  
D 1139 CAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1198  
QY 1240 GGCATGCGGGGCTGAGAGAGAGCGGCTGAGCTGACGTCAGAG 1280  
D 1199 GGGCAGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1239

RESULT 8  
US-09-925-1  
Sequence 1, Application US/09620925  
Patent No. 6468986

GENERAL INFORMATION:  
APPLICANT: Zuckermann et al.

TITLE OF INVENTION: Compositions and Methods for  
Polynucleotide Delivery

NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:

ADDRESSEE: Chiron Corporation  
STREET: 4560 Horton Street  
CITY: Emeryville

STATE: California  
COUNTRY: U.S.A.  
ZIP: 94608-2916

COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk

COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/620,925  
FILING DATE: 21-Jul-2000  
CLASSIFICATION: <unknown>

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/910,647

FILING DATE: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Fujita, Sharon M.  
REGISTRATION NUMBER: 38,459  
REFERENCE/DOCKET NUMBER: 1218,002

TELECOMMUNICATION INFORMATION:  
TELEPHONE: (510) 923-2706  
TELEFAX: (510) 655-3542

INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:

LENGTH: 9600 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)  
SEQUENCE DESCRIPTION: SEQ ID NO: 1:

US-09-620-925-1  
Query Match 4.7%; Score 66.6; DB 4; Length 9600;  
Best local Similarity 45.5%; Pred. No. 1.3e-06;

Matches 237; Conservative 0; Mismatches 284; Indels 0; Gaps 0;

QY 760 GTGCAGGCAATGATCAACCCAGATGAGAGAGTGTGCCAGACATTGAGGACAAACAGCCGC 819  
D 719 GGGCAGGACAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 778  
QY 820 AGACAGCAACTGTTAAACCCAGAGGTTGAGAGCCCTGTCGCGGTTTGGAGAGAGCGC 879  
D 779 GGGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 838  
QY 880 AAGGCGCAACTGCTTCAAGCACTGCCCCGAGCAGAGAGAGAAATTGCAAGCCGCTGCGG 939  
D 839 GAGGGGCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 898  
QY 940 GGCCTCATCCGCCAGTACGAGACCACTTGGAGGCGTCTCAAACTGCTGAGATCCGCC 999  
D 899 CAGGAGCAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 958  
QY 1000 ATCCAGTCCATGAGAGAGCCGACAGTGGCTCTTACCTCCAGAGGCAAAAGAGCTGATC 1059  
D 959 CAGGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1018  
QY 1060 AACAAAGTCCGGGCGCAATGTCGAAGGTGAGCTGCAAGAGCGCCGAGACCCAGGCTATGAG 1119  
D 1019 GAGCAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1078  
QY 1120 AGCATGAGCAATTCCTGTGATGAGCTGAGACACGTGGCCGAATTTTGGAAACCATCGAC 1179  
D 1079 GGGCAGGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1138  
QY 1180 TTCAGCCGGGCGCGCTGGGATGAAGAGATGACGACATGCTTTGGATGGGAGAGAG 1239  
D 1139 CAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1198  
QY 1240 GGCATGCGGGGCTGAGAGAGAGCGGCTGAGCTGACGTCAGAG 1280  
D 1199 GGGCAGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1239

RESULT 9  
US-07-884-811-15  
Sequence 15, Application US/07884811  
Patent No. 5316921

GENERAL INFORMATION:  
APPLICANT: Godowski, Paul J. Lokker, Nathalie A. Mark, Melanie R.

TITLE OF INVENTION: SINGLE-CHAIN HEPATOCYTE GROWTH FACTOR VARIANTS  
NUMBER OF SEQUENCES: 21  
CORRESPONDENCE ADDRESS:

ADDRESSEE: Genentech, Inc.  
STREET: 460 Point San Bruno Blvd  
CITY: South San Francisco

STATE: California  
COUNTRY: USA













ORGANISM: EBNA  
US-09-130-114-2

Query Match 4.5%; Score 64; DB 2; Length 1931;  
Best Local Similarity 44.9%; Pred. No. 3.2e-06;

Matches 284; Conservative 0; Mismatches 345; Indels 3; Gaps 1;

```

OY 653 TCGGCGCCCAAGAGACTGTGAGTGCGCCCTCTGCCCACATTATTAACAGCCGAAGA 712
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1179 TGTGGGCTTACTACTACTGATGATGATGATGATGATGATGATGATGATGATGATG 1120
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 713 GTGAGTGAAGCATTGATGATGATGATGATGATGATGATGATGATGATGATGATG 772
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1119 TGGAGAGGGGAGCCGGAAGAGTGTCTAAGGGGGGAGCCGGAAGATGATGATGATG 1060
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 773 TCACCCAGATGAGAGAGTGTGTCAGACCTTGAAGACAAACAGCCGACAGAGAAGAAC 832
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1059 CGCCGAGAGTGTATGAGAGAGCTGGGGCCGAGGTGATGAGAGAGCTGGGGCGGAGTGA 1000
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 833 TGTTAACCCAGAGGTTCAGAGACCTGTGCGGGTTCAGAGAGCCGAAGGCGAATGC 892
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 999 CGGAGAGAGCTGGGGCCGAGGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 940
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 893 TTCAAGCACTGCGCCCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 952
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 939 ACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 880
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 953 AGTACGAGAGACCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1012
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 879 AGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 820
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1013 AGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1072
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 819 AGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 760
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1073 CAATGTCAAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1132
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 759 ACGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 700
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1133 TCTCTGTAGCGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1192
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 699 AGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 643
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1193 CCGCTGGGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1252
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 642 AGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 583
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1253 TGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1284
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 582 AGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 551
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

# RESULT 17

US-08-232-463-14/c  
Sequence 14, Application US/08232463  
Patent No. 5670367

## GENERAL INFORMATION:

APPLICANT: DORNER, F.  
APPLICANT: SCHEIFLINGER, F.  
APPLICANT: FALKNER, F. G.  
TITLE OF INVENTION: RECOMBINANT FOXLPOX VIRUS  
NUMBER OF SEQUENCES: 52  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Foley & Lardner  
STREET: 1800 Diagonal Road, Suite 500  
CITY: Alexandria  
STATE: VA  
COUNTRY: USA  
ZIP: 22313-0299

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/232.463  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/07/935.313  
FILING DATE:  
APPLICATION NUMBER: EP 91 114 300.6  
FILING DATE: 26-AUG-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: BENT, Stephen A.  
REGISTRATION NUMBER: 29,768  
REFERENCE/DOCKET NUMBER: 30472/114 IMMU  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703)836-9300  
TELEFAX: (703)683-4109  
TELEX: 899149  
INFORMATION FOR SEQ ID NO: 14:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 7218 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
CLONE: pTZgpt-Fls  
US-08-232-463-14

Query Match 4.2%; Score 59.8; DB 1; Length 7218;  
Best Local Similarity 6.0%; Pred. No. 4.7e-05;  
Matches 25; Conservative 224; Mismatches 166; Indels 0; Gaps 0;

```

OY 981 AAGCTGTGTGAGTCCGCCATCCAGTCATGAGAGAGAGAGAGAGAGAGAGAGAGAG 1040
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1455 AGAGATAGAGAAATTTGGTACRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1396
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1041 GCAGCAAGAGAGCTGTCAACAGTCCGGGCAATGCGAAGTGGAGTGGCGAGAGAGC 1100
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1395 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1336
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1101 GCCGAGCAGAGCTATGAGAGATGAGCAATCTCTGTGAGCGTGGAGAGAGAGAGAG 1160
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1335 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1276
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1161 AATGTCCAGACCATGACTTCAGCCGGCGCGCTGGGATGAAGAGATGAGAGACAT 1220
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1275 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1216
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1221 GCGTTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1280
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1215 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1156
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1281 AGGCTCAGGCTGCACTGAGCCGACTCTGATCCAGAGAGAGAGAGAGAGAGAGAGAG 1340
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1155 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1096
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1341 GGGATGCTGAGATGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1395
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1095 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1041
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

## RESULT 18

US-08-728-323A-1  
Sequence 1, Application US/08728323A  
Patent No. 5948676

## GENERAL INFORMATION:

APPLICANT: Chang, Yuan  
APPLICANT: Bohenzky, Roy A.  
APPLICANT: Russo, James J.  
APPLICANT: Edelman, Isidore S.  
APPLICANT: Moore, Patrick S.  
TITLE OF INVENTION: Immediate Early Protein From Kaposi's  
TITLE OF INVENTION: Sarcoma-Associated Herpesvirus, DNA





GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 3, 2002, 09:05:22 ; Search time 64 Seconds  
(without alignments)  
8610.941 Million cell updates/sec

Title: US-09-908-988B-1

Sequence: 1 aagagagtgtagacagagtggtt.....ataaagagctcaagtgctcc 1431

Scoring table: IDENTITY\_NUC  
Gap 10.0, Gapext 1.0

Searched: 341543 seqs, 192557720 residues

Total number of hits satisfying chosen parameters: 683086

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database:

Published Applications NA.\*  
1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:\*  
2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq:\*  
3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq:\*  
4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq:\*  
5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq:\*  
6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq:\*  
7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq:\*  
8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:\*  
9: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq:\*  
10: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq:\*  
11: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq:\*  
12: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq:\*  
13: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq:\*  
14: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq:\*

Pred. NO. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	1431	100.0	1431	10	US-09-908-988B-1 Sequence 1, Appl1
2	484.6	33.9	1597	10	US-09-908-988B-5 Sequence 5, Appl1
3	478.8	33.5	1781	10	US-09-764-864-21 Sequence 21, Appl1
4	450.8	31.5	2580	10	US-09-908-988B-3 Sequence 3, Appl1
5	439.6	30.7	1762	10	US-09-764-864-34 Sequence 34, Appl1
6	229.2	16.0	587	10	US-09-764-864-493 Sequence 493, Appl1
7	162	11.3	446	10	US-09-764-864-482 Sequence 482, Appl1
8	141.6	9.9	1039	10	US-09-880-192-25 Sequence 25, Appl1
9	59.6	4.2	1852	10	US-09-969-852-4 Sequence 4, Appl1
10	50.8	3.5	1400	10	US-09-764-864-556 Sequence 56, Appl1
11	50.4	3.5	3441	10	US-09-880-192-41 Sequence 41, Appl1
12	48	3.4	7065	9	US-09-911-496-115 Sequence 115, Appl1
13	48	3.4	7065	10	US-09-874-923-115 Sequence 115, Appl1
14	47.4	3.3	768	9	US-09-938-842A-812 Sequence 812, Appl1
15	46.4	3.2	280	10	US-09-923-876-5584 Sequence 5584, Appl1
16	45.2	3.2	1969	10	US-09-864-761-3972 Sequence 3972, Appl1
17	43.8	3.1	1464	10	US-09-771-161A-54 Sequence 54, Appl1
18	43.8	3.1	2518	10	US-09-771-161A-53 Sequence 53, Appl1
19	43.8	3.1	659158	9	US-09-771-208-20 Sequence 20, Appl1

20	43.6	3.0	319	10	US-09-960-352-6282 Sequence 6282, Appl1
21	43.6	3.0	1044	10	US-09-969-708-167 Sequence 167, Appl1
22	43.6	3.0	1044	10	US-09-880-107-2241 Sequence 2241, Appl1
23	43.6	3.0	4854	9	US-09-963-875-2 Sequence 2, Appl1
24	43.4	3.0	2108	10	US-09-962-832-225 Sequence 225, Appl1
25	42.8	3.0	765	10	US-09-759-143-524 Sequence 524, Appl1
26	42.8	3.0	765	10	US-09-780-669-524 Sequence 524, Appl1
27	42.8	3.0	765	10	US-09-822-827-524 Sequence 524, Appl1
28	42.8	3.0	1119	10	US-09-759-143-177 Sequence 177, Appl1
29	42.8	3.0	1119	10	US-09-780-669-177 Sequence 177, Appl1
30	42.8	3.0	1119	10	US-09-030-606-177 Sequence 177, Appl1
31	42.8	3.0	1119	10	US-09-822-827-177 Sequence 177, Appl1
32	42.8	3.0	1119	10	US-09-115-453-177 Sequence 177, Appl1
33	42.8	3.0	1214	10	US-09-759-143-225 Sequence 225, Appl1
34	42.8	3.0	1214	10	US-09-780-669-225 Sequence 225, Appl1
35	42.8	3.0	1214	10	US-09-822-827-225 Sequence 225, Appl1
36	42.8	3.0	1214	10	US-09-115-453-326 Sequence 326, Appl1
37	42.8	3.0	1215	10	US-09-759-143-326 Sequence 326, Appl1
38	42.8	3.0	1215	10	US-09-780-669-326 Sequence 326, Appl1
39	42.8	3.0	1215	10	US-09-822-827-326 Sequence 326, Appl1
40	42.8	3.0	1265	10	US-09-759-143-173 Sequence 173, Appl1
41	42.8	3.0	1265	10	US-09-780-669-173 Sequence 173, Appl1
42	42.8	3.0	1265	10	US-09-030-606-173 Sequence 173, Appl1
43	42.8	3.0	1265	10	US-09-822-827-173 Sequence 173, Appl1
44	42.8	3.0	1265	10	US-09-115-453-173 Sequence 173, Appl1
45	42.8	3.0	1350	10	US-09-759-143-616 Sequence 616, Appl1

#### ALIGNMENTS

RESULT 1  
US-09-908-988B-1 Application US/09908988B  
Sequence 1, US20020127690A1  
Patent No. US20020127690A1  
GENERAL INFORMATION:  
APPLICANT: OLSON, ERIC  
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR STABILIZING MICROTUBULES  
FILE REFERENCE: MYOG:02805  
CURRENT APPLICATION NUMBER: US/09/908, 988B  
CURRENT FILING DATE: 2000-07-18  
PRIOR APPLICATION NUMBER: 60/219, 020  
PRIOR FILING DATE: 2000-07-18  
NUMBER OF SEQ ID NOS: 6  
SOFTWARE: Patentin Ver. 2.1  
SEQ ID NO 1  
LENGTH: 1431  
TYPE: DNA  
ORGANISM: Mus musculus  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (199)..(1296)  
US-09-908-988B-1

Query Match 100.0%; Score 1431; DB 10; Length 1431;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1431; Conservative 0; Mismatches 0; Indels 0; Caps 0;

QY	1	AAGAGCTGACACAGAGCTGCTGAATACACAGAGGCTGAGAGAGCTGTTAGGGGAGG 60
DB	1	AAGGAGTCTGACACAGAGCTGCTGAATACACAGAGGCTGAGAGAGCTGTTAGGGGAGG 60
QY	61	GACGAGCTCTTCACAGAGGAGCAATAGCCGGGATCCCAAGATCCAGTCAGCTTAAC 120
DB	61	GACGAGCTCTTCACAGAGGAGCAATAGCCGGGATCCCAAGATCCAGTCAGCTTAAC 120
QY	121	TGACCGAGGAGGAGGTCACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 180
DB	121	TGACCGAGGAGGAGGTCACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 180
QY	181	GCTCCAGAGAGGCGCGGAGATGACTTCACGCTGGGTTTCAAGCGCGCTGCTAGGGGATGCG 240

Db	181	GTCTCAGAGCGCGCGGGATGAACTTCAAGTGGGGTTTAAAGCCGCTGCTAGGGGATGCG	240
QY	241	CACACATGGAACACTTGGAGAGCAGCTCATTTGCCATCTGCTCGTAGAGATGTTCTCC	300
Db	241		
QY	301	AAGCCCGGAGATTTCCCGCCCAACACACCTGTCCGGAAGGTGCGCAAGCACGCTC	360
Db	301		
QY	361	TTCCAGGCTCTAATCTCTGTGGCAATCCGGGGCTCCACAACGGTGTCTTCAGAGGA	420
Db	361		
QY	421	CGTTCCGATCCCATCTGTAGACAGAGTGTCTGTGAAAGGATGTGTCTATGCG	480
Db	421		
QY	481	CTGCAAGGGAACCTGTAGTGGAGAACATCATTTGACATCTACAAACAGAGATCTCCGG	540
Db	481		
QY	541	CCATGCAAGCCAAAGCGTGAACACAGCACTCATGTGTAGAGACAGAGCAGAGAGATC	600
Db	541		
QY	601	AACATCTACTGCTGAGCTGGAGGTGCCACCTCTCTCTGTGAAGGTTTGGCGGC	660
Db	601		
QY	661	CACAAAGACTGTAGGTGGCCCTCTGCCCCACCATTTTCAAAACGCCAGAGAGTGAAGTGTG	720
Db	661		
QY	721	AGGATGGCATCGGATGCTGTGTGGCGGCATGACGCTGTGAGAGGAGTGTATCAACCG	780
Db	721		
QY	781	ATGAGAGAGGTGTGCCAGCCATTGAGAGCAACAGCCGACAGACAAACAACTGTTAAAC	840
Db	781		
QY	841	CAGAGGTTCAGAACCCGTGTGGGGTTTGGAGAGCGCAAGGGCGAATGCTTCAAGCA	900
Db	841		
QY	901	CTGGCCCGGAGCAGAGAGGAAGTTGACAGGCGCTGGGGGCTCATCGGCAGTACGGA	960
Db	901		
QY	961	GACCACTTGGAGGGCTCTTAAAGCTGTGTGATGCCCATCCAGTCCATGTGAGAGACCG	1020
Db	961		
QY	1021	CAGATGCTCTTACTCTCCAGACGGCAAAGAGCGTATCAACAGAGTGGGGCAGATGTGCG	1080
Db	1021		
QY	1081	AAGGTGAGCTGCAAGACGGCCGGAGCCAGGCTATGAGAGCATGTGAGCAATTTCTGTG	1140
Db	1081		
QY	1141	AGCGTGGAGCACTGTGCGCGAAATGTTCCGAACCATGTGACTTCCAGCCGGGGCCGCTGGG	1200
Db	1141		
QY	1201	GATTGAAGAGATGACACATGTGCTTGGATGGGAGAGAGGGCAATGCGGGCTTGAGAGAG	1260
Db	1201		
QY	1261	GAGCGGCTGAGGTGCCAGAAAGGCTCAGGCTGACAGCCCACTTGATCAGAGCGC	1320
Db	1261		

Db	1261	GAGCGCTGGACGTGCGCAGAGGCTCAGGCCCTGACCTGACCCACTCTGTATCCAGAGCCG	1320
QY	1321	ACACCCGGAAGCGGAGACCAAGAGATGCTGAGATCTGCGCAGAGACACACCGGCCACCA	1380
Db	1321	ACACCCGGAAGCGGAGACCAAGAGATGCTGAGATCTGCGCAGAGACACACCGGCCACCA	1380
QY	1381	GCTCGGCTTCCGCCCCCGGGAAGGTTCTCATTAAGAGACTCAAGTGTCCC	1431
Db	1381	GCTCGGCTTCCGCCCCCGGGAAGGTTCTCATTAAGAGACTCAAGTGTCCC	1431
RESULT 2			
US-09-908-988B-5			
; Sequence 5, Application US/09908988B			
; Patent No. US20020127690A1			
; GENERAL INFORMATION:			
; APPLICANT: OLSON, ERIC			
; APPLICANT: SPENCER, JEFFREY A.			
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR STABILIZING MICROTUBULES			
; TITLE OF INVENTION: IN STRIPED MUSCLE CELLS			
; FILE REFERENCE: MTOG:02805			
; CURRENT APPLICATION NUMBER: US/09/908, 988B			
; CURRENT FILING DATE: 2000-07-18			
; PRIOR APPLICATION NUMBER: 60/219, 020			
; PRIOR FILING DATE: 2000-07-18			
; NUMBER OF SEQ ID NOS: 6			
; SOFTWARE: PatentIn Ver. 2.1			
; SEQ ID NO 5			
; LENGTH: 1597			
; TYPE: DNA			
; ORGANISM: Mus musculus			
; FEATURE:			
; NAME/KEY: CDS			
; LOCATION: (299)..(1327)			
US-09-908-988B-5			
Query Match			
Best Local Similarity 33.9%; Score 484.6; DB 10; Length 1597;			
Matches 677; Conservative 69.2%; Pred. No. 2.6e-107;			
Matches 677; Conservative 0; Mismatches 299; Indels 3; Gaps 1;			
QY	226	CTGCTAGGGGATGCGCCACAACATGAGACAACCTTGGAGAAAGCACTCATTTGCCCATCTTGC	285
Db	317	CTGATTTCTGTANAGAAACGTAATTTGGAGAACCTGTATCTGCCCCATCTGC	376
QY	286	CTGAGATGTTCTCCAAACCCGCTGTGATCTTGCCCTGCAACACAACTGTGCGCGAAG	345
Db	377	CTGAGATGTTTACCAACCCGTGTGATCTTGCCCTGCAACACAACTGTGCGCGAAG	436
QY	346	TGTGCCAACGACGCTTCCAGGCTCTTAATCTCTGTGGCAATCCCGGGCTCCACAAG	405
Db	437	TGTGCCAACGACATCTTCCAGGCTGCGAATCTCTACTGACCAACCGCGGTGCTCAGTG	496
QY	406	GTCGCTTAGAGAGACGTTTCCGATGCCCATCTGTAGGCAACGAGTGTCTCGACAGG	465
Db	497	TTCATGTCGTGAGAGGTCGTTTCCGTTGCCCTGTGCGCCGCAATGATATCTGACCGG	556
QY	466	CATGCTGTATGAGCTGTGACGCGAAGCTGTAGTGAAGAACATCATTTGACATCTACAAG	525
Db	557	CACGGGGGTGTACGGCTGTGACGAGAACCTGCTGTGGAAACATCATTTGACATCTACAAG	616
QY	526	CAGAGTCTCTCCCGGCACTGCAAGCCGCAAGGCTGAACAGCACTCATTTGTGAGAGCAC	585
Db	617	CAGAGTGTCTCAG---TCGGCCCTGTGCAAGAAAGGCAAGCCACCGATGTGCAAGGAAACAC	673
QY	586	GAGAGAGAGAAGATCAACATCTACTGCTGTGAGTGTGAGAGTGCACCACTGTCTCTCTGC	645
Db	674	GAAAGACAGAAGATCAACATCTACTGCTGTGAGTGTGAGAGTGTGAGAGTGTCTCTCTGC	733
QY	646	AAGGTTTTCGGCGCCACAGAGACTGTGAGTGTGCGCCCTCTGCCACCATTTACAACGC	705
Db	734	AAGGTTTTCGGCGCCACAGAGACTGTGAGTGTGCGCCCTCTGCCACCATTTACAACGA	793
QY	706	CAGAAAGTGTGAGCTGAGCAGATGAGCATGCGCATGTGTGAGGGGGCAATGACCTGTGTAG	765

```

Db      794 CAGAGAGCTAGAGCTGAGTAAGCTGATCTGATCTCATGCTGTGGTGGGGAACGACCGAGTGCAG 853
Qy      766 GCAGTGAATCCACCCAGATGGAGAGGTGTGCCAGACCATTTGAGGACAACAGCCGACAGACG 825
Db      854 ACATATCTCTCTAGCTGGAGAGGAGCTGTGCAGAGTATGACCAAGAGAAATPAGCCACAGAGTG 913
Qy      826 AAGCACTGTTAAACCAAGAGAGTTGAGACCCCTGTGGCGGCTTTTGGAGAGAGCCGACAGGCG 885
Db      914 AAGGAGAGGTGTGAGTCAAGAGTTTGGACACCTCTTACGCTCCATCTTGATGAGAAAGAGCG 973
Qy      886 GAACGTGTTAAGACACTGAGCCCGGAGACAGAGAGAGAAATGTGACGCCGTGCGGGGCTTC 945
Db      974 GAGCTGTGGACGGGAGATCAGCAGCAAGACAGGAGGAAAGAGTGGAGTGCATGAGAGGCTGTG 1033
Qy      946 ATCCGCCAGTACGAGAGACCCACTTGGAGGGCTCTTCAAAAGCTGGGTGGAGTCCGCCATCCAG 1005
Db      1034 ATCTCCAGATACAGGGAGCGAGCTGGAAAAGTCCACAAAGCTTTGGAGAGACCGCATCCAG 1093
Qy      1006 TCCATGAGAGAGCGCCGAGATGGCTCTCTTACTCTCCAGCAGAGCAAGAGAGCTGATCAACAG 1065
Db      1094 TCCCTGATGAGACCCCGGAGGGGCTTACTTCTCTCAGTGGCCAAAGCACTGATCAAGAGC 1153
Qy      1066 GTCCGGGCAATGTGGAAGGTGGAGCTGGCAGGAGCGCCGAGCCAGGCTATGAGACATG 1125
Db      1154 ATTGTAAAGGCTCTCAAGGGGCTGCCAGCTGGGGAAGACAGACGAAGGCTTTGAGAACATG 1213
Qy      1126 GAGCAATTCCTGTGAGCTGTGAGCAGTGCAGCTGGGCCCAAAATGTTGCAACCATGCACTTCAG 1185
Db      1214 GAGTACTTACTGTGAGCTTGAACACATATGACAGAGGCTTGGAGGCTTATGAGCTTGGG 1273
Qy      1186 CCGGGCGCCGCTGGGGATG 1204
Db      1274 ACAGGTAAAGGATGTGATG 1292

RESULT 3
US-09-764-864-21
: Sequence 21, Application US/09764864
: Patent No. US20020132753A1
: GENERAL INFORMATION:
: APPLICANT: Rosen et al.
: TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
: FILE REFERENCE: PT23
: CURRENT APPLICATION NUMBER: US/09/764,864
: PRIORITY FILING DATE: 2001-01-17
: Prior application data removed - consult PALM or file wrapper
: NUMBER OF SEQ ID NOS: 1792
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 21
: LENGTH: 1781
: TYPE: DNA
: ORGANISM: Homo sapiens
US-09-764-864-21

Query Match      33.5%; Score 478.8; DB 10; Length 1781;
Best Local Similarity 66.4%; Pred. No. 6.7e-106;
Matches 703; Conservative 0; Mismatches 352; Indels 3; Gaps 1;

Qy      226 CTGCTAGGGGATGCGCACCAACATGAGACAATGGAGAGCAGCTCATTTGCCCATCTGC 285
Db      137 CTGATCCAGAGATGGGAATCCATGAGAGAACTTGGAGAAAGCACTGATGCGCTATFCTGC 196
Qy      286 CTGAGATGTTCTCCAGAGCCGTGTGATCTTGGCCCTGCCAACAACCTGTGCCGCAAG 345
Db      197 CTGGAGATGTTTACCAACGACAGTGTGTCTTCTTCCGTGCACACAACTGTGCGCGAAG 256
Qy      346 TGTGCACAGACGCTTCCAGAGCGCTCTATCTCTGTGGCAATTCGCGGGGCTCCACAAG 405
Db      257 TGTGCATATGATATTTCCAGGCTCTCAAAATCCCTTACTGAGACCAAGCCGGGCAAGCTCAGTG 316
Qy      406 GTGTCTTAGAGAGAGCTTTCCGATGCCCATCTTGTAGGACAGAGAGTTTCTCTGACAG 465

```

Db	317	TCATGTCGTGGAGGCCGCTTTCCGCTGGCCCAACCTGCGCCGACGAGGTGATCATGATCGT	376
QY	466	CATGCTGTCTATGCGCTTCGACGCGAACCCTGCTAGTGGAGAAACATCATTTGACATCTACAG	525
Db	377	CACGAGATTTTACGGCCCTGCGACAGGAACTGCTGGTGGAGAAACATCATCGACATCTACAAA	436
QY	526	CAGGAGTCCTCCCGGCCACTGCAGCCCAAGGCTGCACAGCACTCATGTGTGAGGAGCAC	585
Db	437	CAGAGTGTCTCTCACTCGGCCGCTGCAGAAAGGCGAGT ---CACCCCATGTGCAAGGAGCAC	493
QY	586	GAGGACGAGAAATCATCATCTACTGCTTGAGCTGCGAGTGGCCCACTGCTCTCTCTGC	645
Db	494	GAAATGTGAAAAAATAACATCTACTGTCCTCAAGTGTGAGTGGCCCACTGCTCATGTGCG	553
QY	646	AAGTTTTCGGCGGCCCAAGAGACGTGTAGGTGGGCCCCCTGCGCCACATTTTACAAACG	705
Db	554	AAGTGTTTGGGATTCACAAAGGCTCTGCAGAGTGGCCCCATTGTCAAGTGTCTTCCAGGGA	613
QY	706	CAGAAAGTGTGAGCTGAGCGGATGGCATCGCATGCTGTGTGGCGGGCATGACCGTGTGAG	765
Db	614	CAAAAGACTGGAACCTGAATAACTGTATCTCCATMGCTGTGGCGGGGAAGACCGTGTGAG	673
QY	766	GCACGTATCACCAGANTGTGAGAGAGGTGTGCCAGACCATTTGAGGAAACAAGCCGACAGAC	825
Db	674	ACCATCATCTCATCTGTGAGAGATTCCTCGTAGTGTACCAAGGAAACAGTACCAGGTA	733
QY	826	AAGCAACTGTTAAACCAAGGCTTCAGACCCCTGTGTCGGGTGTTTGGAGAGCGCAAGGCG	885
Db	734	AAGGAAGAGCTGTAGCGACAAAGTTTACACGTTGTATGTGCATCTGTGATGAGAAGAAAGT	793
QY	886	GAACCTGTTCAAGCACTGGCCCGGGAGCAGAGAGCAAACTTGCAGCGCTGCGGGGCTC	945
Db	794	GAGTGTCTGCAGCGGATCACGAGGAGCAGAGAAAGAGCTTAAGTATGAGAGGCCCTC	853
QY	946	ATCCGCCAGTACGAGACCACTTGGAGGGCTCTCAAAAGCTGTGGAGTTCGCCCATTCAG	1005
Db	854	ATTCAGACGTACCAAGAGCAGCTGTGACAAATTCACAAAGCTGTGTGAATCTCCATCCAG	913
QY	1006	TCCATGAGAGACCCCGCATGTGGCTTCTAACCCTCAGCAGGCAAGGAGAGCTGTACAACAG	1065
Db	914	TCCCTGAGACAGCCTGGGGGAGCCACTTCTCTTGACGTGCCCAACCACTCATCAAAAGC	973
QY	1066	GTCCGGGCGCATGTCCAAAGTGTGAGCTGTGCAGAGACGGCCGGAGCCAGGGCTATGAGAGCATG	1125
Db	974	ATTGTGTGAAGCTTCCAAAGGCGTCCAGCTGTGGGAAAGACAGACAGGAGCTTTGAGAACTG	1033
QY	1126	GAGCAATTCCTGTGAGCGTGTGAGCAGCTGTGGCCGAATTTTGGCAACCATCATCTTCAG	1185
Db	1034	GACTTCTTACTTGTGATTTAAGACCAATAGCAGACGCCCTGAGAGCCATTGACTTTGGG	1093
QY	1186	CCGGCCGCCGCTGGGGATGAAAGATGACGACATGCTTGTGGATGGGAGAGAGGGCAAT	1245
Db	1094	ACAGATGAGGAAGAGCAAAATTCATTTGTAAGAAAGAAATCAGAGGAAGAGAAAGTCCACA	1153
QY	1246	GGGGGCTGTGAGGAGAGCGGCTGTGACGTGCGCAGAGG 1283	
Db	1154	GAAAGGAAAGAAAGAGACCAACGATGAGGAGCTGTGATG 1191	
RESULT 4			
US-09-908-988B-3			
; Sequence 3, Application US/09908988B			
; Patent No. US20020127690A1			
; GENERAL INFORMATION:			
; APPLICANT: OLSON, ERIC			
; APPLICANT: SPENCER, JEFFREY A.			
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR STABILIZING MICROTUBULES			
; TITLE OF INVENTION: IN STRIPED MUSCLE CELLS			
; FILE REFERENCE: MYOG:02805			
; CURRENT APPLICATION NUMBER: US/09/908, 988B			
; CURRENT FILING DATE: 2000-07-18			
; PRIOR APPLICATION NUMBER: 60/219, 020			
; PRIOR FILING DATE: 2000-07-18			











SEQ ID NO 56  
 LENGTH: 1400  
 TYPE: DNA  
 ORGANISM: Homo sapiens  
 US-09-764-864-56

Query Match 3.5%; Score 50.8; DB 10; Length 1400;  
 Best Local Similarity 51.3%; Pred. No. 0.0033;  
 Matches 118; Conservative 0; Mismatches 112; Indels 0; Gaps 0;

559 GAACAGACACCTGATGTGTAGAGACAGAGAGAGAGAGATCAACATCTACTGCTGAGC 618  
 440 GACCCGAGACCCAGAGTGTGCGTGCACACCGAGACCCGCTCTTCTTGGGAGAG 499  
 619 TCGGAGGTCCCACTGCTCTCTCTGCAAGTTTGGCCGCCCAAGAGACTGTGAGTG 678  
 500 GACCCAGAGCTCATCTGTGCGCTGCGTCTGCTGCGCTCCCAACACACCCGCTC 559  
 679 GCGCCCTGCTCCCACTGCTCTCTCTGCAAGTTTGGCCGCCCAAGAGACTGTGAGTG 738  
 560 AGCCCGCTCTCCACCTGCTGCGACCCGATAGAGAGAGAGCTCGACGCCCTCTTCTGAG 619  
 739 CTGCTGCGCGGCAATGACCGTGTGACAGGAGTATCACCCAGATGAGAGA 788  
 620 CTGAACGAGAGAGAGAGAGAGTGTGATGACCTCATGCCCAACTGTGTGA 669

## RESULT 11

US-09-880-192-41  
 Sequence 41, Application US/09880192  
 Patent No. US20020077470A1  
 GENERAL INFORMATION:  
 APPLICANT: Walker, Michael G.  
 APPLICANT: Volkmut, Wayne  
 APPLICANT: Klingler, Tod M.  
 APPLICANT: Azimzai, Yalda  
 TITLE OF INVENTION: POLYNUCLEOTIDES ASSOCIATED WITH CARDIAC MUSCLE FUNCTION  
 FILE REFERENCE: PB-0009-1 CIP  
 CURRENT APPLICATION NUMBER: US/09/880,192  
 CURRENT FILING DATE: 2001-06-12  
 NUMBER OF SEQ ID NOS: 62  
 SOFTWARE: PERL Program  
 SEQ ID NO 41  
 LENGTH: 3441  
 TYPE: DNA  
 ORGANISM: Homo sapiens  
 FEATURE:  
 NAME/KEY: misc.feature  
 OTHER INFORMATION: Incyte ID No. US20020077470A1 4151935CB1  
 US-09-880-192-41

Query Match 3.5%; Score 50.4; DB 10; Length 3441;  
 Best Local Similarity 44.3%; Pred. No. 0.0055;  
 Matches 204; Conservative 0; Mismatches 256; Indels 0; Gaps 0;

578 AGAGAGACGAGAGAGAGATCAACATCTACTGCTGAGCTGCGAGGTGCCACCTGCT 637  
 1088 AAGAGCTGAAAAGTCCAGATTGACATCTGTACACTGCAAAATGCTCAATTTCTG 1147  
 638 CTCTGCAAGTTTGGGCGCCCAAGAGACTGTGAGTGGCCCTCTGCGCCACATTT 697  
 1148 CCACTGACAGGTGTGTGGCACCCCAAGACCATGAAGTTTCAACGCTTGACACAGTA 1207  
 698 ACAAGCCCAAGAGAGTGTGAGCATGAGCATGCGATGCTGCTGCTGCGGCGCATGACC 757  
 1208 TAAGTGTGTAAAGTTCAATTAGCAGAAATTTAGAAAATTTACAAAAGTCTCTGA 1267  
 758 GTGTGACGAGATGATACCATGAGAGAGTGTGCGACCATGTGAGACACAGCC 817  
 1268 GAGTTGAGAGCTTTGTAGTATGATAGATATCTTTTAAATACATTTGAGGAAACTGTA 1327  
 818 GCAGACAGAACTGTAAACAGAGTTGAGAGCCCTGTGCGGCTTTTGGAGAGAC 877

Db 1328 GTAAATATGAGAAAGCTAGAGAAAGATGAGAAATGATGAAAGTTTACAC 1387  
 QY 878 GCAGGGGGAATGCTTCAAGACACTGGCCGGAGAGAGAGAGTGTGACAGCGCTG 937  
 Db 1388 AGTATGATGAGAAAGCCAGAGCTTTGAGAAATGAGAGAGAGATGAGATCTCTG 1447  
 QY 938 GGGGCTCATCCGCAAGTACAGAGACCACTGTGAGGGCTCTCAAGCTGTGTGAGTCC 997  
 Db 1448 ATGACACATGTGCTCACTTTTGCAGAGCATGAGACCTGCCAAAGACACCTTGAGACCA 1507  
 QY 998 CCATCCAGTCCATGAGAGAGCCGAGATGCTCTACT 1037  
 Db 1508 TCGTAGAGAAACAGAGAGAGCTTGTATGAGGCGCTTCTCT 1547

## RESULT 12

US-09-991-496-115  
 Sequence 115, Application US/09991496  
 Patent No. US20020169285A1  
 GENERAL INFORMATION:  
 APPLICANT: Reed, Steven G.  
 APPLICANT: Campos-Neto, Antonio  
 APPLICANT: Webb, John R.  
 APPLICANT: Dillon, David C.  
 APPLICANT: Skelky, Yasir A.W.  
 APPLICANT: Bhatia, Ajay  
 APPLICANT: Colier, Rhea  
 APPLICANT: Probst, Peter  
 APPLICANT: Brannon, Mark  
 TITLE OF INVENTION: LEISHMANIA ANTIGENS FOR USE IN THE  
 TITLE OF INVENTION: THERAPY AND DIAGNOSIS OF LEISHMANIASIS  
 FILE REFERENCE: 210121.420C9  
 CURRENT APPLICATION NUMBER: US/09/991,496  
 CURRENT FILING DATE: 2001-11-20  
 NUMBER OF SEQ ID NOS: 137  
 SOFTWARE: FastSeq for Windows Version 4.0  
 SEQ ID NO 115  
 LENGTH: 7065  
 TYPE: DNA  
 ORGANISM: Leishmania major and chagasi  
 US-09-991-496-115

Query Match 3.4%; Score 48; DB 9; Length 7065;  
 Best Local Similarity 46.9%; Pred. No. 0.026;  
 Matches 254; Conservative 0; Mismatches 280; Indels 8; Gaps 3;

QY 846 GTTCGAGACCTGTGCGGCTTTTGGAGAGCGGCAAGGGGAGACTGCTTACAGACTGCG 905  
 Db 2190 GCTGAGAGGACAGAGTGTGCGGCTGTGCGCGGAGACCGGAGAGGCGGCGAGCTGGC 2249  
 QY 906 CCGGAGAGAGAGAGAGTGTGAGCGGCTGTGCGGCGCTCATCCGAGTACGAGACCA 965  
 Db 2250 CGGAGAGCGGAGAGAGTGTGAGCGGCTGTGAGCGGCTGTGAGCGGCTGTGAGCGG 2309  
 QY 966 CTGTGAGGCTCTCTCAAGTGTGTGAGTCCGCGCATCACTC--CATGAGAGAGCGCA 1022  
 Db 2310 GAGAGCAAGTTGACAGCGCTGTGCGCGGAGACCGGAGAGGCGGCGAGCTGGCGC 2369  
 QY 1023 GATGCTCTTACTCTCAGAGAGCAAGAGATGATCAACAAGTGTGAGGCGCAATGTGAA 1082  
 Db 2370 GAAGCGGAGAGAGTGTGAGAGCGCTGTGAGCGGCTGTGAGCGGCTGTGAGCGGCG 2429  
 QY 1083 GGTGAGCTGTGAGAGAGCGGCGGAGAGCTGTGAGAGCATGAGCAATCTGTGAG 1142  
 Db 2430 GGCACAGTGTGACAGCTGTGCGCGGAGACCGGAGAGGCGCGGCGAGCACTGGCGGCA 2489  
 QY 1143 CGTGAGACAGTGTGCGGCAATGTTTGCAGACCATGCTTCCAGCGCGGCGCGCTGGGA 1202  
 Db 2490 CGCGAGAGAGCTG--CAGAGCGCTAGACACCGCACGAGAGAGCGCGCGGAGCTGGA 2546  
 QY 1203 TGAAGAGATGAGACACTGTGCTTTGATGAGAGAGAGGCAATGCGGAGCGGAGAGGA 1262  
 Db 2547 GGCACAGTGTGACAGCTGTGCGCGGAGACCGGAGAGGCGCGGAGCACTGGCGCGCA 2606





```
; GENERAL INFORMATION:
; APPLICANT: LEVINE, et al.
; TITLE OF INVENTION: VARIANTS OF PROTEIN KINASES
; FILE REFERENCE: 802620-2005.1
; CURRENT APPLICATION NUMBER: US/09/771,161A
; PRIOR FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: 09/724,676
; PRIOR FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: 136776
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: 135619
; PRIOR FILING DATE: 2000-04-12
; NUMBER OF SEQ ID NOS: 273
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 54
; LENGTH: 1464
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: -
; LOCATION: (1)-(1464)
; OTHER INFORMATION: "n" can be any nucleotide 'a', 'c', 'g' or 't'
US-09-771-161A-54
```

```
Query Match          3.1%; Score 43.8; DB 10; Length 1464;
Best Local Similarity 44.7%; Pred. No. 0.16;
Matches 263; Conservative 0; Mismatches 317; Indels 9; Gaps 2;
```

```
OY 442 AGGACGAGGTTGCTCTGGACGAGCATGTGTATGCGCTGCAGCGGACCTGCTAGTG 501
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 460 AGCCGCGTGGCGGCTGGAGAGAGCATGTCATGAGTGAGAGTGAACAGGGGGCG 519

OY 502 GAGAACATCATTTGACATCTACAGAGAGTCCCTCCGGCCACTGCAGCCCAAGCTGAA 561
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 520 GAGAACATGATCCAGACCTACAGCATGCGACAGCAAGAGCCGGAAGCTGCTGACA 579

OY 562 CAGCACTCATGTGTGAGAGAGCAGAGAGAGAGATCAATCTACTGCTGAGCTGC 621
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 580 GCCCAGCAGATGTGGAGAGGACAGTAAAGCAAGTATGACATCATCCGATCCGATCCGC 639

OY 622 GAGTGCCCACTGCTCTCTCTGCAAGGTTTTGGCGCCCAAGAGACTGTGAGTGCC 681
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 640 CGGGGCGTCGAGCGCCAGCAGCTGAGAACCAAGGCAAGCCCGGATGACACCCAAAGGAGT 699

OY 682 CCTGTGCCCAACATTTACAAAGCCAGAAAGTGAAGTGAAGTGGATGGATGGATGCTG 741
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 700 CTTGACCTTGGGGGCTGTGAGCTGCCATGAAAGAGCTGGCGCACCACTTCCGAGTGAG 759

OY 742 GTGGCGGGAATGACCGTGTGAGGAGAGTATGACCGAGATGAGAGAGT---GTGCCAG 798
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 760 CACGGGTTGGCGGAGGTCGCAAGAGTACTGGCCTGCTCAAGCGCTGCCAAGGCCCG 819

OY 799 ACCATTGAGGACAAGCCGCGAGAGAGCAACTGTTA-----AACAGAGGTTGAG 852
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 820 GACCCGAAGGAGCTGACGAGGCGCAGAGAAATTGACAATCCAAACCAAGAGCTGGGG 879

OY 853 ACCCTGTGCGGGTTTGGAGAGGAGGAGGGAAGTGTTCAGACATGAGCTGGCCCGGAG 912
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 880 CTGCTGGGGAAGCTGTGAGGCGGAGACTTGGGAGCTGGCCCGCGACACCCCAAGGGG 939

OY 913 CAGGAGAGAAAGTTGAGCGGCTGCGGGGCTCATCCGCAAGTACGAGAGACCACTTGAG 972
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 940 CGGCTGCTCGAAGAGAGCTCGCTGGGGCTCTCTCCGCTGCTTCAAGACCCGCTGGCC 999

OY 973 GGCTCTCAAGAGCTGAGAGTCCGCCATCAGTCCATGAGAGAGCCGC 1021
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1000 GGGCCTTTCGCCGACGACTACAGCAACCTGTGCAAGCCCGCGCCGC 1048
```

```
RESULT 18
US-09-771-161A-53
; Sequence 53, Application US/09771161A
; Patent No. US2002011081A1
```

```
; GENERAL INFORMATION:
; APPLICANT: LEVINE, et al.
; TITLE OF INVENTION: VARIANTS OF PROTEIN KINASES
; FILE REFERENCE: 802620-2005.1
; CURRENT APPLICATION NUMBER: US/09/771,161A
; PRIOR FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: 09/724,676
; PRIOR FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: 136776
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: 135619
; PRIOR FILING DATE: 2000-04-12
; NUMBER OF SEQ ID NOS: 273
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 53
; LENGTH: 2518
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: -
; LOCATION: (1)-(2518)
; OTHER INFORMATION: "n" can be any nucleotide 'a', 'c', 'g' or 't'
US-09-771-161A-53
```

```
Query Match          3.1%; Score 43.8; DB 10; Length 2518;
Best Local Similarity 44.7%; Pred. No. 0.15;
Matches 263; Conservative 0; Mismatches 317; Indels 9; Gaps 2;
```

```
OY 442 AGGACGAGGTTGCTCTGGACGAGCATGTGTATGCGCTGCAGCGGACCTGCTAGTG 501
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 460 AGCCGCGTGGCGGCTGGAGAGAGCATGTCATGAGTGAGAGTGAACAGGGGGCG 519

OY 502 GAGAACATCATTTGACATCTACAGAGAGTCCCTCCGGCCACTGCAGCCCAAGCTGAA 561
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 520 GAGAACATGATCCAGACCTACAGCATGCGACAGCAAGAGCCGGAAGCTGCTGACA 579

OY 562 CAGCACTCATGTGTGAGAGAGCAGAGAGAGAGATCAATCTACTGCTGAGCTGC 621
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 580 GCCCAGCAGATGTGGAGAGGACAGTAAAGCAAGTATGACATCATCCGATGCAACTCCGC 639

OY 622 GAGTGCCCACTGCTCTCTCTGCAAGGTTTTGGCGCCCAAGAGACTGTGAGTGCC 681
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 640 CGGGGCGTGCAGCGCCAGCACTGAGAGAACAGAGAGCGCCGAGATGACACCAAGGAGT 699

OY 682 CCTGTGCCCAACATTTACAAAGCCAGAAAGTGAAGTGAAGTGGATGGATGGATGCTG 741
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 700 CTTGACCTTGGGGGCTGTGAGCTGCCATCGAAGAGCTGGCGCACCACTTCCGAGTGAG 759

OY 742 GTGGCGGGAATGACCGTGTGAGGAGAGTATGACCGAGATGAGAGAGT---GTGCCAG 798
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 760 CACGGGTTGGCGGAGGTCGCAAGAGTACTGGCCTGCTCAAGCGCTGCCAAGGCCCG 819

OY 799 ACCATTGAGGACAAGCCGCGAGAGAGCAAGCAACTGTTA-----AACAGAGGTTGAG 852
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 820 GACCCGAAGGAGTACAGCAGAGGCCAGAGAGAAATTGACAAGATCCAAAGAGCTGGGG 879

OY 853 ACCCTGTGCGGGTTTGGAGAGGCGCAAGGGCGAAGTCTTCAAGCACTGAGCCCGGAG 912
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 880 CTGCTGGGGAAGCTGTGAGGCGGAGACTTGGGAGCTGGCCCGCGACCAACCCAAAGGGG 939

OY 913 CAGGAGAGAAAGTTGAGCGGCTGCGGGGCTCATCCGCAAGTACGAGAGACCACTTGAG 972
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 940 CGGCTGTGAGAGAGAGCTCGCTGGGGCTCTCTCCGCTGCTTCAAGACCCGCTGGCC 999

OY 973 GGCTCTCAAGAGCTGAGAGTCCGCCATCAGTCCATGAGAGAGCCGC 1021
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1000 GGGCCTTTCGCCGACGACTACAGCAACCTGTGCAAGCCCGCGCCGC 1048
```

```
RESULT 19
US-09-771-208-20/c
; Sequence 20, Application US/09771208
; Patent No. US2002015564A1
; GENERAL INFORMATION:
; APPLICANT: MEDRANO, JUAN
; APPLICANT: BRADFORD, ERIC
; APPLICANT: HORVAT, SIMON
```

```

; TITLE OF INVENTION: CLONING OF A HIGH-GROWTH GENE
; FILE REFERENCE: 407T-9237100S
; CURRENT APPLICATION NUMBER: US/09/771.208
; CURRENT FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: US 08/999,477
; PRIOR FILING DATE: 1997-12-29
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 20
; LENGTH: 659158
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (123459)..(123478)
; OTHER INFORMATION: n is unidentified a, c, g, or t
; NAME/KEY: misc_feature
; LOCATION: (602466)..(602485)
; OTHER INFORMATION: n is unidentified a, c, g, or t
; NAME/KEY: misc_feature
; LOCATION: (494715)..(494814)
; OTHER INFORMATION: n is unidentified a, c, g, or t
; NAME/KEY: misc_feature
; LOCATION: (390986)..(391005)
; OTHER INFORMATION: n is unidentified a, c, g, or t
; NAME/KEY: misc_feature
; LOCATION: (346860)..(346823)
; OTHER INFORMATION: n is unidentified a, c, g, or t
; NAME/KEY: misc_feature
; LOCATION: (317174)..(317193)
; OTHER INFORMATION: n is unidentified a, c, g, or t
; NAME/KEY: misc_feature
; LOCATION: (280353)..(280373)
; OTHER INFORMATION: n is unidentified a, c, g, or t
; NAME/KEY: misc_feature
; LOCATION: (271829)..(271848)
; OTHER INFORMATION: n is unidentified a, c, g, or t
; NAME/KEY: misc_feature
; LOCATION: (183872)..(183891)
; OTHER INFORMATION: n is unidentified a, c, g, or t
; NAME/KEY: misc_feature
; LOCATION: (170625)..(170645)
; OTHER INFORMATION: n is unidentified a, c, g, or t
; NAME/KEY: misc_feature
; LOCATION: (132680)..(132700)
; OTHER INFORMATION: n is unidentified a, c, g, or t
; NAME/KEY: misc_feature
; OTHER INFORMATION: n is a, c, g, or t
; US-09-771-208-20

Query Match      3.1%; Score 43.8; DB 9; Length 659158;
Best Local Similarity 49.8%; Pred. No. 1.1;
Matches 111; Conservative 0; Mismatches 112; Indels 0; Gaps 0;

QY 1040 AGCAGCAAGAGAGCTGATCAACAAGTCCGGGCAATGTCGAAGGTGAGCTGCAGAC 1099
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 251104 AGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 251045

QY 1100 GCGCGAGAGCCAGGCTATGAGAGCATGAGCAATTCTGTGAGCGTGGAGCAGCGTGGCCG 1159
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 251044 AGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 250985

QY 1160 AAATGTTGCGAACCATGACTTCCAGCGGGCGCGCTGGGGATGAAGATGACAGACA 1219
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 250984 AGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 250925

QY 1220 TGGCTTTGATGGGAGAGAGGCAATGCGGGGCTGGAGAGAGA 1262
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 250924 AAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGA 250882
```

```

RESULT 20
US-09-960-352-6282/c
; Sequence 6282, Application US/09960352
; Patent No. US20020137139A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Nengbing
; APPLICANT: Byatt, John C.
; APPLICANT: Mathiasagan, Nagappan
; TITLE OF INVENTION: NOCETIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; FILE REFERENCE: 16511.006/37-21(10298)C
; CURRENT APPLICATION NUMBER: US/09/960.352
; CURRENT FILING DATE: 2001-09-24
; NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 6282
; LENGTH: 319
; TYPE: DNA
; ORGANISM: Bos taurus
; OTHER INFORMATION: Clone ID: 27-LIB3058-020-Q1-K1-G3
; US-09-960-352-6282

Query Match      3.0%; Score 43.6; DB 10; Length 319;
Best Local Similarity 50.0%; Pred. No. 0.11;
Matches 109; Conservative 0; Mismatches 109; Indels 0; Gaps 0;

QY 1086 GGAAGCTGGCAGAGAGCGCCGAGCCAGGCTATGAGAGCATGGAGCAATTCTGTGAGCGT 1145
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 283 GCGCAGAGCCGAGAGGCGCGGCAAGATGAGAGCGCACTATCCAGAGATCGAGAGCGCGCC 224

QY 1146 GGAAGCAGTGGCCGCAAAATGTTGCGAACCATGACATTCAGCGCGGCGCGCTGGGATGA 1205
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 223 GGAAGCCCGCGGCGGCGGCTGTGAGGCGCCAGGCGCGCGCGCGCCGCGCGAGAGATGC 164

QY 1206 AGAGATGACAGACATGGCTTTGATGGGAGAGAGGCAATGCGGGGCTGGAGAGAGAGCG 1265
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 163 TGCCGTAGCGGCCGAGCAGAGAGAGAGGCGGCTAGGCGCAAGGCCAGAGCGCCAGAGAG 104

QY 1266 GCTGGAGAGTGCAGAAAGGCTATGAGGCTGACACTGACCCG 1303
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 103 GCGGAGAGGCGCCGAGGCGCTCCCGCAGAGGACGAGCAG 66
```

Search completed: December 3, 2002, 10:57:29  
Job time : 1350 secs





GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 3, 2002, 10:09:37 ; Search time 2857 Seconds

(Without alignments)  
14576.865 Million cell updates/sec

Title: US-09-908-988b-1

Sequence: 1431  
1 aagagagtagacagagagtg.....ataaagagactcaagtgctccc 1431

Scoring table: OLIGO\_NUC  
Gapop 60.0 , Gapext 60.0

Searched: 2054640 seqs, 14551402878 residues

Word size: 13

Total number of hits satisfying chosen parameters: 36437

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database:

Genbank:\*

1: gb\_ba:\*

2: gb\_ba:\*

3: gb\_in:\*

4: gb\_om:\*

5: gb\_ov:\*

6: gb\_pat:\*

7: gb\_ph:\*

8: gb\_pl:\*

9: gb\_pr:\*

10: gb\_ro:\*

11: gb\_sts:\*

12: gb\_sy:\*

13: gb\_un:\*

14: gb\_vi:\*

15: em\_ba:\*

16: em\_fun:\*

17: em\_hum:\*

18: em\_in:\*

19: em\_mu:\*

20: em\_om:\*

21: em\_or:\*

22: em\_ov:\*

23: em\_pat:\*

24: em\_ph:\*

25: em\_pl:\*

26: em\_ro:\*

27: em\_sts:\*

28: em\_un:\*

29: em\_vi:\*

30: em\_hum:\*

31: em\_hum:\*

32: em\_hum:\*

33: em\_hum:\*

34: em\_hum:\*

35: em\_hum:\*

36: em\_hum:\*

37: em\_hum:\*

38: em\_hum:\*

39: em\_hum:\*

40: em\_hum:\*

41: em\_hum:\*

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1431	100.0	1431	6 AX418848	AX418848 Sequence
2	1431	100.0	1448	10 AF294790	AF294790 Mus muscu
3	367	25.6	209211	2 AC084883	AC084883 Mus muscu
4	236	16.5	199296	2 AC114619	AC114619 Mus muscu
5	173	12.1	49743	2 AC114541	AC114541 Mus muscu
6	57	4.0	164772	2 AC095184	AC095184 Rattus no
7	36	2.5	1329	2 HSA291714	HSA291714 Homo sapi
8	36	2.5	118847	9 AC013413	AC013413 Homo sapi
9	35	2.4	1500	6 AX060632	AX060632 Sequence
10	29	2.0	123280	2 AC025662	AC025662 Homo sapi
11	28	2.0	1597	6 AX418852	AX418852 Sequence
12	28	2.0	196461	10 AL627314	AL627314 Mouse DNA
13	27	1.9	1746	9 AB047601	AB047601 Macaca fa
14	27	1.9	1750	9 BC007750	BC007750 Homo sapi
15	27	1.9	1810	9 HSA243488	HSA243488 Homo sapi
16	27	1.9	1925	9 AK091728	AK091728 Homo sapi
17	27	1.9	2098	9 HSA243489	HSA243489 Homo sapi
18	27	1.9	2202	9 HSA291712	HSA291712 Homo sapi
19	27	1.9	2634	9 AK091310	AK091310 Homo sapi
20	27	1.9	121721	9 AC090195	AC090195 Homo sapi
21	27	1.9	183912	2 AC021240	AC021240 Homo sapi
22	22	1.5	3372	1 AF250140	AF250140 Pseudomon
23	22	1.5	153005	2 AC011976	AC011976 Homo sapi
24	22	1.5	157839	2 AC015718	AC015718 Homo sapi
25	22	1.5	188205	2 AC073385	AC073385 Homo sapi
26	22	1.5	203375	9 AC083841	AC083841 Homo sapi
27	21	1.5	91428	9 AC079902	AC079902 Homo sapi
28	21	1.5	96299	9 AC006330	AC006330 Homo sapi
29	21	1.5	108690	2 AC107289	AC107289 Rattus no
30	21	1.5	148994	2 AC128011	AC128011 Rattus no
31	21	1.5	159510	2 AC102387	AC102387 Mus muscu
32	21	1.5	180759	2 AC093136	AC093136 Pan trogl
33	21	1.5	184444	2 AC025065	AC025065 Homo sapi
34	21	1.5	192341	2 AC107613	AC107613 Homo sapi
35	21	1.5	237364	2 AC122289	AC122289 Homo sapi
36	20	1.4	606	9 HSA36987	HSA36987 Mus muscu
37	20	1.4	608	9 HS305G6	HS305G6 H. sapiens
38	20	1.4	10029	1 AE010904	AE010904 Methanosa
39	20	1.4	10821	2 AC014410	AC014410 Drosophill
40	20	1.4	16173	3 AF262977	AF262977 Drosophill
41	20	1.4	29952	2 AC012715	AC012715 Drosophill
42	20	1.4	40834	7 BA1242593	BA1242593 Bacteriop
43	20	1.4	72305	8 NCBI3A5	NCBI3A5 Neurospor
44	20	1.4	74317	8 NC12F11	NC12F11 Neurospor
45	20	1.4	87463	8 OSJN00225	OSJN00225 Oryza sat

## ALIGNMENTS

RESULT 1  
AX418848  
LOCUS  
DEFINITION  
AX418848  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
house mouse.  
Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
Sequence 1 from Patent WO0206318.  
AX418848.1 GI:21523712

DNA linear PAT 18-JUN-2002

REFERENCE  
AUTHORS  
TITLE  
Olson, E.N. and Spencer, J.A.  
Methods and compositions for stabilizing microtubules and  
intermediate filaments in striated muscle cells

Pred. No. is the number of results predicted by chance to have a

JOURNAL Patent: WO 020618-A 1 24-JAN-2002;  
BOARD OF REGENTS THE UNIVERSITY OF TEXAS SYSTEM (US)  
FEATURES Location/Qualifiers  
source 1. 1431  
/organism="Mus musculus"  
/db\_xref="taxon:10090"  
199. .1299  
/note="unnamed protein product"  
/codon\_start=1  
/protein\_id="CAD35455.1"  
/db\_xref="GI:21523713"

CDS

DEDDMDALDGEBSNAGLEERLDVPEBSGLH"  
ATVLOAKELINKYKAMSKYELAGRPERYSEMEPFSVSEVHAEMLRITIDRPGAS  
TICVALEERKELLQALAREOEKLRQYRGLIRQYDHLKESKRYESAIOSEEPOM  
APLPTIYKROKSELSDGILAMLVAGNDRVQAVITOMEVQOITEDNSRROKOLNORE  
NIIIDYKQESSRPLHAKAEQHMCEHEDEKINIVCLSEVPTCSIKYFGAHKCEV  
MLCRKANDVFOASNPLOWSRSTVSSGSRFCRSHREVVLDHRYVGLORNLVE  
/translation="MNFVTFKPLLDGHNMDNEKOLICPIELEMFKPVVILPCOH  
MLCRKANDVFOASNPLOWSRSTVSSGSRFCRSHREVVLDHRYVGLORNLVE  
NIIIDYKQESSRPLHAKAEQHMCEHEDEKINIVCLSEVPTCSIKYFGAHKCEV  
APLPTIYKROKSELSDGILAMLVAGNDRVQAVITOMEVQOITEDNSRROKOLNORE  
TICVALEERKELLQALAREOEKLRQYRGLIRQYDHLKESKRYESAIOSEEPOM  
ATVLOAKELINKYKAMSKYELAGRPERYSEMEPFSVSEVHAEMLRITIDRPGAS  
DEDDMDALDGEBSNAGLEERLDVPEBSGLH"

BASE COUNT 338 a 384 c 473 g 236 t  
ORIGIN

Query Match 100.0%; Score 1431; DB 6; Length 1431;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1431; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 AAGAGTGTAGACAGAGTGTCTGGAATATAGACAGGGGTAGAGAGACTGTAGGGAGAG 60  
1 AAGAGGTGTAGACAGAGTGTCTGGAATATAGACAGGGGTAGAGAGACTGTAGGGAGAG 60  
61 GACAGGACTCTTCCAAAGAGGAGCAATAGCCGGGATCCCAAGATTCAGTCAAGCTTAAC 120  
61 GACAGGACTCTTCCAAAGAGGAGCAATAGCCGGGATCCCAAGATTCAGTCAAGCTTAAC 120  
121 TGACCGAGAGAGGTGACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 180  
121 TGACCGAGAGAGGTGACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 180  
121 TGACCGAGAGAGGTGACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 180  
181 GCTCCAGAGAGCGCGGGGAGTGAAGTTCACGGGTTCACAGCCGCTGCTAGGGATGCG 240  
181 GCTCCAGAGAGCGCGGGGAGTGAAGTTCACGGGTTCACAGCCGCTGCTAGGGATGCG 240  
181 GCTCCAGAGAGCGCGGGGAGTGAAGTTCACGGGTTCACAGCCGCTGCTAGGGATGCG 240  
181 GCTCCAGAGAGCGCGGGGAGTGAAGTTCACGGGTTCACAGCCGCTGCTAGGGATGCG 240  
241 CACACATGAGCAACTTGGAGAGAGAGTCAATGCCCCCATTTGCTGGAGATGTTCTCC 300  
241 CACACATGAGCAACTTGGAGAGAGAGTCAATGCCCCCATTTGCTGGAGATGTTCTCC 300  
301 AAGCCCGTGTGATCTTGCCCTGACCAACACACTGTGCGGCAAGTGTGCCAACGACGTC 360  
301 AAGCCCGTGTGATCTTGCCCTGACCAACACACTGTGCGGCAAGTGTGCCAACGACGTC 360  
361 TTCAGAGCCCTCTAATCTCTGTGTGCAATCCCGGGGCTCCACAGCGTGTCTTACAGAGA 420  
361 TTCAGAGCCCTCTAATCTCTGTGTGCAATCCCGGGGCTCCACAGCGTGTCTTACAGAGA 420  
421 CGTTCCGATGCCATCTTGTAGGACAGAGGTGTCTGAGACAGGAGCATGTCTATATGGC 480  
421 CGTTCCGATGCCATCTTGTAGGACAGAGGTGTCTGAGACAGGAGCATGTCTATATGGC 480  
481 CTGACAGCGAAGCTGTAGTGAAGACATCTTGCATCTCAACACAGAGAGCTCTCCGG 540  
481 CTGACAGCGAAGCTGTAGTGAAGACATCTTGCATCTCAACACAGAGAGCTCTCCGG 540  
481 CTGACAGCGAAGCTGTAGTGAAGACATCTTGCATCTCAACACAGAGAGCTCTCCGG 540  
481 CTGACAGCGAAGCTGTAGTGAAGACATCTTGCATCTCAACACAGAGAGCTCTCCGG 540  
541 CCAGCTGACGCCAAGCTGTAGACACACTCATGTGTGAGAGACAGAGAGAGAGATC 600  
541 CCAGCTGACGCCAAGCTGTAGACACACTCATGTGTGAGAGACAGAGAGAGAGATC 600  
601 AACATCTACGCTGAGCTGAGAGTGGCCGACCTCTCTCTGCAAGGTTTTCGGCGCC 660  
601 AACATCTACGCTGAGCTGAGAGTGGCCGACCTCTCTCTGCAAGGTTTTCGGCGCC 660  
661 CACAGAGAGTGTAGAGTGGCCCTCTGCGCCACCATTTTACAAAGCCAGAGAGTGAAGT 720  
661 CACAGAGAGTGTAGAGTGGCCCTCTGCGCCACCATTTTACAAAGCCAGAGAGTGAAGT 720

721 AGCGATGGCATCGCATGCTGTGTGCGGGCAATGACCGTGTGCGAGCATGATCACCAG 780  
721 AGCGATGGCATCGCATGCTGTGTGCGGGCAATGACCGTGTGCGAGCATGATCACCAG 780  
781 ATGGAGAGAGTGTGCCAGACCATTTAGAGACACAGCCGACAGAGAGCAACTGTAAAC 840  
781 ATGGAGAGAGTGTGCCAGACCATTTAGAGACACAGCCGACAGAGAGCAACTGTAAAC 840  
841 CAGAGGTTGCAAGCCCTGTGCGGGCTTTTGGAGAGAGGAGGAGGAGGAGGAGGAGGAG 900  
841 CAGAGGTTGCAAGCCCTGTGCGGGCTTTTGGAGAGAGGAGGAGGAGGAGGAGGAGGAG 900  
901 CTGGCCCGGAGACAGAGAGAGAGAGTGTGACAGCGGTGCGGGGCTCATCCGCAATACGA 960  
901 CTGGCCCGGAGACAGAGAGAGAGAGTGTGACAGCGGTGCGGGGCTCATCCGCAATACGA 960  
961 GACCACTTGGAGAGGCTCTTCAAAAGCTGTGTGAGTCCGCCATCCAGTCCATGAGAGAGCG 1020  
961 GACCACTTGGAGAGGCTCTTCAAAAGCTGTGTGAGTCCGCCATCCAGTCCATGAGAGAGCG 1020  
1021 CAGATGGCTCTTACCTCCAGAGCAAGAGAGCTGATCAACAAAGTTCGGGCAATGTCG 1080  
1021 CAGATGGCTCTTACCTCCAGAGCAAGAGAGCTGATCAACAAAGTTCGGGCAATGTCG 1080  
1081 AAGGTGAGCTGTGACAGAGAGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1140  
1081 AAGGTGAGCTGTGACAGAGAGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1140  
1141 AGCGTGGAGACAGTGTGCGGCAATTTGCGACCATGTGACTTCCAGCCGGGCGCGCTGGG 1200  
1141 AGCGTGGAGACAGTGTGCGGCAATTTGCGACCATGTGACTTCCAGCCGGGCGCGCTGGG 1200  
1201 GATGAAGAGATGACACATGCTGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1260  
1201 GATGAAGAGATGACACATGCTGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1260  
1261 GAGCGCTGTGAGAGTGTGCGGCAATTTGCGACCATGTGACTTCCAGCCGGGCGCGCTGGG 1320  
1261 GAGCGCTGTGAGAGTGTGCGGCAATTTGCGACCATGTGACTTCCAGCCGGGCGCGCTGGG 1320  
1321 ACACCGGAAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1380  
1321 ACACCGGAAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1380  
1381 GCTGCGCTTCCCGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1431  
1381 GCTGCGCTTCCCGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1431

RESULT 2  
AF294790 1448 bp mRNA linear ROD 08-SEP-2000  
LOCUS  
DEFINITION Mus musculus RING-finger protein MURF mRNA, complete cds.  
ACCESSION AF294790  
VERSION AF294790.1 GI:9945009  
KEYWORDS  
SOURCE  
ORGANISM Mus musculus.  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 1448)  
Spencer, J.A., Eliazzer, S., Ilarria, R.L. Jr., Richardson, J.A. and  
Olson, E.N.  
TITLE Regulation of microtubule dynamics and myogenic differentiation by  
MURF, a striated muscle RING-finger protein  
JOURNAL J. Cell Biol. 150 (4), 771-784 (2000)  
MEDLINE 20411220  
PUBMED 10953002

REFERENCE 2 (bases 1 to 1448)  
AUTHORS Spencer, J.A. and Olson, E.N.  
REFERENCE Direct Submission  
TITLE Submitted (09-AUG-2000) Molecular Biology, UTSW, 6000 Harry Hines  
Bldv., Dallas, TX 75390-9148, USA

FEATURES  
source Location/Qualifiers  
1..1448  
/organism="Mus musculus"  
/db\_xref="taxon:10090"  
/feature\_type="cardiac muscle"  
199..1299  
/codon\_start=1  
/product="RING-finger protein MURF"  
/protein\_id="AA03076.1"  
/db\_xref="GI:9945010"  
/translation="MNFVGRPLIGDAHNMNLEKILICPLEMFKPVVILPCQH  
NLCKRCAVDYQASNPIMWOSRGSTVSSGSRPCSRREVLDHGVGLDRNLIVE  
NIIIDYKQESSRPLHAKOHLMECEHEDENIKINICSEVPTCSLKYFGAKCOEY  
APLPYIKRQKSELDGIFAMLVAGNDRAVITOMEVQOTEDNSRROKOLLNOFE  
TLCAVLEERKELLQALROBEKIORVGLIROYGDHLESSKIVESAIGMEPROM  
ALYIQAKELINKYKAMSKVELAGRPPEYSEMEDFVSVEHVAEMLRITIDPOGAG  
DEEDDMALDEENAGLEERLDVPESSGLH"

BASE COUNT 355 a 384 c 473 g 236 t  
ORIGIN

Query Match 100.0%; Score 1431; DB 10; Length 1448;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1431; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAGAGGTAGACAGAGTGTGCAAAATAGACAGAGGAGTGTAGAGACTGTAGGGAGAG 60  
DB 1 AAGAGGTAGACAGAGTGTGCAAAATAGACAGAGGAGTGTAGAGACTGTAGGGAGAG 60  
QY 61 GACAGAGCTTTCCAAAGAGAGAGCAATAGCCGGGATCCCAAGATCCAGTCACTTAAC 120  
DB 61 GACAGAGCTTTCCAAAGAGAGAGCAATAGCCGGGATCCCAAGATCCAGTCACTTAAC 120  
QY 121 TGACCGAGGAAGGGGTGCACAGGAGGGGAGGAGCCCAAGAGAGGGGCCACAGGGAGAG 180  
DB 121 TGACCGAGGAAGGGGTGCACAGGAGGGGAGGAGCCCAAGAGAGGGGCCACAGGGAGAG 180  
QY 181 GCTCCAGAGCGCGCGGGGTGTAAGTTCACGAGTGGGTTTCAAGCCGCTGTAGGGAGTGC 240  
DB 181 GCTCCAGAGCGCGCGGGGTGTAAGTTCACGAGTGGGTTTCAAGCCGCTGTAGGGAGTGC 240  
QY 241 CACAACATGAGACACTTGGAGAGAGCACTATTTGCCCATCTGCTGAGATGTTCTCC 300  
DB 241 CACAACATGAGACACTTGGAGAGAGCACTATTTGCCCATCTGCTGAGATGTTCTCC 300  
QY 301 AAGCCGGGTGATGTCGTCGCAACAACACTGTGCGGAGAGTGTGCAAGAGAGTGC 360  
DB 301 AAGCCGGGTGATGTCGTCGCAACAACACTGTGCGGAGAGTGTGCAAGAGAGTGC 360  
QY 361 TTCCAGGCTCTAATCTCTGTGCAATCCCGGGCTCCACAAGGAGTGTTCAGAGAGA 420  
DB 361 TTCCAGGCTCTAATCTCTGTGCAATCCCGGGCTCCACAAGGAGTGTTCAGAGAGA 420  
QY 421 CGTTTCCGATGCCATCTTGTAGGACAGAGTGTCTGTGACAGAGGATGTTCTATGGC 480  
DB 421 CGTTTCCGATGCCATCTTGTAGGACAGAGTGTCTGTGACAGAGGATGTTCTATGGC 480  
QY 481 CTGACGGGAACCTGATGAGAGAAATGATGATGATGATGATGATGATGATGATGATG 540  
DB 481 CTGACGGGAACCTGATGAGAGAAATGATGATGATGATGATGATGATGATGATGATG 540  
QY 541 CCACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 600  
DB 541 CCACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 600  
QY 601 AACATCTACTGCTGAGTGCAGGAGTGCACCTGCTCTCTGCAAGGTTTTCGGCGCC 660  
DB 601 AACATCTACTGCTGAGTGCAGGAGTGCACCTGCTCTCTGCAAGGTTTTCGGCGCC 660  
QY 661 CACAGAGAGTGTAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 720  
DB 661 CACAGAGAGTGTAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 720  
QY 721 AGCGATGCGATGCGATGCTGTGTGGCGGCAATGACCGTGTGCGAGCGATGATCCAG 780

DB 721 AGCGATGCGATGCGATGCTGTGTGGCGGCAATGACCGTGTGCGAGCGATGATCCAG 780  
QY 781 ATGAGAGAGTGTGCGCAGACCATTTAGAGACAACAGCCGACAGACAGACTGTTAAAC 840  
DB 781 ATGAGAGAGTGTGCGCAGACCATTTAGAGACAACAGCCGACAGACAGACTGTTAAAC 840  
QY 841 CAGAGGTTCGAGACCTTGTGGCGGTTTGTAGAGAGCCCAAGGGGCACTGCTTCAAGA 900  
DB 841 CAGAGGTTCGAGACCTTGTGGCGGTTTGTAGAGAGCCCAAGGGGCACTGCTTCAAGA 900  
QY 901 CTGCGCGGGAGCAGAGAGAGAGAGAGTGTGAGCGCGTGTGCGGGGCTCATCCGCAAGCA 960  
DB 901 CTGCGCGGGAGCAGAGAGAGAGAGTGTGAGCGCGTGTGCGGGGCTCATCCGCAAGCA 960  
QY 961 GACCACTTGTGAGAGGCTCTCTCAAGAGTGTGAGTCCGCGCATCATGATGAGAGAGCCG 1020  
DB 961 GACCACTTGTGAGAGGCTCTCTCAAGAGTGTGAGTCCGCGCATCATGATGAGAGAGCCG 1020  
QY 1021 CAGATGCTCTTACCTCTCAGAGAGCAAGAGAGTGTATCAACAGAGTGTGAGAGAGTGTG 1080  
DB 1021 CAGATGCTCTTACCTCTCAGAGAGCAAGAGAGTGTATCAACAGAGTGTGAGAGAGTGTG 1080  
QY 1081 AAGGTGAGCTGTGAGAGAGAGCGCGGAGCGCTATGAGAGCATGAGCAATTTCTGTG 1140  
DB 1081 AAGGTGAGCTGTGAGAGAGAGCGCGGAGCGCTATGAGAGCATGAGCAATTTCTGTG 1140  
QY 1141 AGCGTGAAGCACTGTGCGGCAATTTGCGAGACATGACTTCCAGCGGGCGCGCTGTGG 1200  
DB 1141 AGCGTGAAGCACTGTGCGGCAATTTGCGAGACATGACTTCCAGCGGGCGCGCTGTGG 1200  
QY 1201 GATGAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1260  
DB 1201 GATGAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1260  
QY 1261 GAGCGGCTGTGAGAGTGTGCAAGAGCTGTGACTGACCTGACCTGACCTGATCCAGAGCGC 1320  
DB 1261 GAGCGGCTGTGAGAGTGTGCAAGAGCTGTGACTGACCTGACCTGATCCAGAGCGC 1320  
QY 1321 ACACCCGAAGGGGAGCCCAAGGATGCTGAGATGCTGCCAAGAGCAACCGCGCCCA 1380  
DB 1321 ACACCCGAAGGGGAGCCCAAGGATGCTGAGATGCTGCCAAGAGCAACCGCGCCCA 1380  
QY 1381 GCTGCGCTTCCCGCGCGGGAAGTTCATTAAGAGCACTAAGTGTCC 1431  
DB 1381 GCTGCGCTTCCCGCGCGGGAAGTTCATTAAGAGCACTAAGTGTCC 1431

RESULT 3  
AC084883  
LOCUS  
DEFINITION Mus musculus chromosome 5 clone RP23-154K6 strain C57BL/6J, WORKING  
DRAFT SEQUENCE, 21 unordered pieces.  
ACCESSION AC084883  
VERSION AC084883.10 GI:17976433  
KEYWORDS HTG; HTGS; PHASE1; HTGS\_DRAFT.  
SOURCE Mus musculus.  
ORGANISM Mus musculus.  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 209211)  
Han, J., Montgomery, K.T., Grills, G., Chiu, D., Decker, J., Fusina, M.,  
Goltz, J., Haider, A., Hall, L., Ioshikhes, I.P., Lee, E., Long, J.,  
Perera, A., Shim, C., Thomas, E. and Kuchelapatti, R.  
High Throughput Mouse Sequencing  
Unpublished  
2 (bases 1 to 209211)  
Han, J., Montgomery, K.T., Grills, G., Chiu, D., Decker, J., Fusina, M.,  
Goltz, J., Haider, A., Hall, L., Ioshikhes, I.P., Lee, E., Long, J.,  
Perera, A., Shim, C., Thomas, E. and Kuchelapatti, R.  
Direct Submission  
Submitted (29-Nov-2000) Department of Molecular Genetics, Albert  
Einstein College of Medicine Center, 1300 Morris Park Ave.,

## COMMENT

Bronx, NY 10461, USA  
On Dec 21, 2001 this sequence version replaced g1:15148085.

-----Genome Center

Center: Harvard Partners Genome Center

Center Code: HPGC

Web site: <http://www.hpgc.org/Sequence/mouse.html>  
Contact: [hpgc@medel.mgh.harvard.edu](mailto:hpgc@medel.mgh.harvard.edu)

-----Summary Statistics

Center project name: ABD

Sequencing vector: pUC18: 108752

Chemistry: Dye-terminator Big Dye: 100%

\*Consensus quality: 204023 at least Q20

\*Consensus quality: 203326 at least Q30

\*Consensus quality: 201766 at least Q40

\*Estimated insert size: agarose-FP - N/A

\*Estimated insert size: 208811 - sum-of-contigs

Quality coverage: agarose-FP - N/A

Quality coverage: 10.2 x in Q20 bases; sum-of-contigs estimation

-----  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 21 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

```
1 30155: contig of 30155 bp in length
* 30156 30175: gap of unknown length
* 30176 76428: contig of 46253 bp in length
* 76429 76448: gap of unknown length
* 76449 99269: contig of 22821 bp in length
* 99270 99289: gap of unknown length
* 99290 119029: contig of 19740 bp in length
* 119030 119049: gap of unknown length
* 119050 133795: contig of 14746 bp in length
* 133796 133815: gap of unknown length
* 133816 147516: contig of 13701 bp in length
* 147517 147536: gap of unknown length
* 147537 161508: contig of 13972 bp in length
* 161509 161528: gap of unknown length
* 161529 171555: contig of 10027 bp in length
* 171556 171575: gap of unknown length
* 171576 182235: contig of 10660 bp in length
* 182236 182255: gap of unknown length
* 182256 190163: contig of 7908 bp in length
* 190164 190183: gap of unknown length
* 190184 194456: contig of 4273 bp in length
* 194457 194476: gap of unknown length
* 194477 199061: contig of 4585 bp in length
* 199062 199081: gap of unknown length
* 199082 201315: contig of 2234 bp in length
* 201316 201335: gap of unknown length
* 201336 201781: contig of 446 bp in length
* 201782 201801: gap of unknown length
* 201802 203287: contig of 1486 bp in length
* 203288 203307: gap of unknown length
* 203308 204141: contig of 834 bp in length
* 204142 204161: gap of unknown length
* 204162 204246: contig of 85 bp in length
* 204247 204266: gap of unknown length
* 204267 205695: contig of 1429 bp in length
* 205696 205715: gap of unknown length
* 205716 206923: contig of 1208 bp in length
* 206924 206943: gap of unknown length
* 206944 207841: contig of 898 bp in length
* 207842 207861: gap of unknown length
* 207862 209211: contig of 1350 bp in length.
```

## FEATURES

## source

1. 209211  
/strain="Mus musculus"  
/organism="C57BL/6J"  
/db\_xref="taxon:10090"

```
/chromosome="5"  
/clone="RP23-154K6"  
/sex="male"  
1. 30155  
/note="assembly_name:Contig39"  
30176. 76428  
/note="assembly_name:Contig38"  
76449. 99269  
/note="assembly_name:Contig37"  
99290. 119029  
/note="assembly_name:Contig36"  
119050. 133795  
/note="assembly_name:Contig35"  
clone_end:17  
vector_side:right"  
133816. 147516  
/note="assembly_name:Contig34"  
147537. 161508  
/note="assembly_name:Contig33"  
161529. 171555  
/note="assembly_name:Contig32  
clone_end:SP6  
vector_side:right"  
171576. 182235  
/note="assembly_name:Contig31"  
182256. 190163  
/note="assembly_name:Contig30"  
190184. 194456  
/note="assembly_name:Contig29"  
194477. 199061  
/note="assembly_name:Contig28"  
199082. 201315  
/note="assembly_name:Contig27"  
201336. 201781  
/note="assembly_name:Contig26"  
201802. 203287  
/note="assembly_name:Contig25"  
203308. 204141  
/note="assembly_name:Contig24"  
204162. 204246  
/note="assembly_name:Contig23"  
204267. 205695  
/note="assembly_name:Contig22"  
205716. 206923  
/note="assembly_name:Contig21"  
206944. 207841  
/note="assembly_name:Contig20"  
207862. 209211  
/note="assembly_name:Contig19"  
BASE COUNT 55539 a 49732 c 50169 g 53288 t 483 others  
ORIGIN  
Query Match 25.68 Score 367; DB 2; Length 209211;  
Best Local Similarity 100.0%; Pred. No. 5; 6e-192;  
Matches 367; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 AAGGAGTGTAGACAGAGTGTCTGGAATATAGACAGGGTGAGAGAGAGGCTTAGGGGAAG 60  
|||||  
Db 190769 AAGGAGTGTAGACAGAGTGTCTGGAATATAGACAGGGTGAGAGAGAGGCTTAGGGGAAG 190828  
|||||  
QY 61 GACAGACTCTTCCAGAGGAGCAATAGCCGGGATCCCAAGATCCAGTACGCTTAAC 120  
|||||  
Db 190829 GACAGACTCTTCCAGAGGAGCAATAGCCGGGATCCCAAGATCCAGTACGCTTAAC 190888  
|||||  
QY 121 TGACCGAGGAAGGTTCCACAGGAGGGAAGAAGCCACAGCAGGCGCACAGGAGG 180  
|||||  
Db 190889 TGACCGAGGAAGGTTCCACAGGAGGGAAGAAGCCACAGCAGGCGCACAGGAGG 190948  
|||||  
QY 181 GCTCCAGAGCGCGCGGATGACTTCACGCTGGGTTTCAAGCGGCTGATGGGATCG 240  
|||||  
Db 190949 GCTCCAGAGCGCGCGGATGACTTCACGCTGGGTTTCAAGCGGCTGATGGGATCG 191008  
|||||  
QY 241 CACACATGAGCAACTTGGAGAGAGCAGCTCATTTGCCCATCTGCTGGAGATGTTTCC 300
```

OY	191009	CACAACTGAGCAACTTGGAGAAGCAGCTCATTTGGCCCATCGTCGTGAGATGTCGCC	191068
OY	301	AAGCCCGTGATCTTGGCCCTGCACACACACCTGTGCCGAAGTGTGCCACGAGCTC	360
Dbl	191069	AAGCCCGTGATCTTGGCCCTGCACACACACCTGTGCCGAAGTGTGCCACGAGCTC	191128
OY	361	TTCGACG 367 TTTTTTT	
Dbl	191129	TTCGACG 191135	
RESULT 4			
LOCUS	AC114619		
DEFINITION	Mus musculus clone RP24-86123, WORKING DRAFT SEQUENCE, 15 ordered pieces.	199296 bp DNA linear	HTG 06-JUN-2001
ACCESSION	AC114619		
VERSION	AC114619.2	GI:21327612	
KEYWORDS	HTG; HTGS_PHASE2; HTGS_DRAFT; HTGS_FULLTOP.		
SOURCE	house mouse.		
ORGANISM	Mus musculus		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus		
AUTHORS	Birren,B., Linton,L., Nisbaum,C. and Lander,E. Mus musculus, clone RP24-86123 Unpublished 2 (bases 1 to 199296)		
JOURNAL			
TITLE			
REFERENCE	Birren,B., Linton,L., Nisbaum,C., Lander,E., All,A., Allen,N., Anderson,S., Barra,N., Bastien,V., Bloom,T., Bocuslavsky,L., Boukhgalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collimore,A., Cook,A., Cooke,P., Dearellano,K., Dewar,R., Diaz,J.S., Dodge,S., Fato,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S., Glade,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N., Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A., Karatis,A., Kells,C., LaRoque,K., Lamazares,R., Landers,T., Lehoczyk,J., Levine,R., Lindblad-Toh,K., Liu,G., Maclean,C., MacDonald,P., Major,J., Margulis,N., Matthews,C., McCarthy,M., McEwan,P., McKernan,K., Meldrum,J., Menues,L., Mihova,T., Mienga,V., Murphy,T., Naylor,T., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Raymond,C., Retta,R., Riebeck,M., Riley,R., Rise,C., Rogov,P., Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupack,R., Seaman,S., Severly,P., Spencer,B., Stange-Thomann,N., Stojanovic,N., Strauss,N., Subramanian,A., Talamas,U., Tesfaye,S., Theodore,J., Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliou,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J., Zemdek,L., Zimmer,A. and Zody,M. Direct Submission Submitted (10-MAR-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA 3 (bases 1 to 199296)		
JOURNAL			
TITLE			
REFERENCE			
AUTHORS	Birren,B., Linton,L., Nisbaum,C., Lander,E., All,A., Allen,N., Anderson,S., Barra,N., Bastien,V., Bloom,T., Bocuslavsky,L., Boukhgalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collimore,A., Cook,A., Cooke,P., Dearellano,K., Dewar,R., Diaz,J.S., Dodge,S., Fato,S., Ferreira,P., Fitzgerald,M., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S., Glade,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N., Grand-Pierre,N., Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A., Karatis,A., Kells,C., LaRoque,K., Lamazares,R., Landers,T., Lehoczyk,J., Levine,R., Lindblad-Toh,K., Liu,G., Maclean,C., MacDonald,P., Major,J., Margulis,N., Matthews,C., McCarthy,M., McEwan,P., McKernan,K., Meldrum,J., Menues,L., Mihova,T., Mienga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Raymond,C., Retta,R., Riebeck,M., Riley,R., Rise,C., Rogov,P., Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupack,R., Seaman,S., Severly,P., Spencer,B., Stange-Thomann,N., Stojanovic,N., Strauss,N., Subramanian,A., Talamas,U., Tesfaye,S., Theodore,J., Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliou,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J., Zemdek,L., Zimmer,A. and Zody,M. Direct Submission Submitted (10-MAR-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA 3 (bases 1 to 199296)		

TITLE  
 JOURNAL  
 COMMENT

Stojanovic, N., Strauss, N., Subramanian, A., Talamas, J., Testa, S.,  
 Theodorou, J., Topham, K., Travers, M., Travis, N., Trigg, J.,  
 Vassiliou, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J.,  
 Young, G., Zainoun, O., Zemdek, L., Zimmer, A. and Zody, M.  
 Direct Submission  
 Submitted (06-JUN-2002) Whitehead Institute/MIT Center for Genome  
 Research, 320 Charles Street, Cambridge, MA 02141, USA  
 On Jun 6, 2002 this sequence version replaced g1:193111148.  
 All repeats were identified using RepeatMasker:  
 Smt, A.F.A. & Green, P. (1996-1997)  
 http://ftp.genome.washington.edu/RM/RepeatMasker.html

-----  
 Genome Center  
 Center: Whitehead Institute/ MIT Center for Genome Research  
 Center code: WtBR  
 Web site: http://www-seq.wi.mit.edu  
 Contact: sequence\_submissions@genome.wi.mit.edu  
 -----  
 Project Information  
 Center project name: L24325  
 Center clone name: 86.L123

-----  
 Summary Statistics  
 Sequencing vector: Plasmid; n/a; 100% of reads  
 Chemistry: Dye-terminator Big Dye; 100% of reads  
 Assembly program: Phrap; version 0.960731  
 Consensus quality: 193147 bases at least Q40  
 Consensus quality: 196185 bases at least Q30  
 Consensus quality: 197247 bases at least Q20  
 Insert size: 194000; agarose-1p  
 Insert size: 197896; sum-of-contigs  
 Quality coverage: 6.1 in Q20 bases; agarose-1p  
 Quality coverage: 6.0 in Q20 bases; sum-of-contigs

-----  
 \* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 15 contigs. Gaps between the contigs  
 \* are represented as runs of N. The order of the pieces  
 \* is believed to be correct as given, however the sizes  
 \* of the gaps between them are based on estimates that have  
 \* provided by the submittor.  
 \* This sequence will be replaced  
 \* by the finished sequence as soon as it is available and  
 \* the accession number will be preserved.

	1	728: contig of 728 bp in length
*	729	828: gap of 100 bp
*	829	2029: contig of 1201 bp in length
*	2030	2129: gap of 100 bp
*	2130	3803: contig of 1674 bp in length
*	3804	3903: gap of 100 bp
*	3904	6051: contig of 2148 bp in length
*	6052	6151: gap of 100 bp
*	6152	10082: contig of 3931 bp in length
*	10083	10182: gap of 100 bp
*	10183	18316: contig of 8134 bp in length
*	18317	18416: gap of 100 bp
*	18417	27114: contig of 8698 bp in length
*	27115	27214: gap of 100 bp
*	27215	33934: contig of 6720 bp in length
*	33935	34034: gap of 100 bp
*	34035	44466: contig of 10432 bp in length
*	44467	44566: gap of 100 bp
*	44567	59489: contig of 14923 bp in length
*	59490	59589: gap of 100 bp
*	59590	79430: contig of 19841 bp in length
*	79431	79530: gap of 100 bp
*	79531	99545: contig of 20015 bp in length
*	99546	99645: gap of 100 bp
*	99646	123290: contig of 23645 bp in length
*	123291	123390: gap of 100 bp
*	123391	151687: contig of 28297 bp in length
*	151688	151787: gap of 100 bp
*	151788	199296: contig of 47509 bp in length.

Location/Qualifiers  
 1. 199296  
 /organism="Mus musculus"  
 /db\_xref="taxon:10090"



10412 11144: contig of 733 bp in length  
11145 11244: gap of 100 bp  
11245 11929: contig of 685 bp in length  
11930 12029: gap of 100 bp  
12030 12728: contig of 700 bp in length  
12730 12829: gap of 100 bp  
12830 13553: contig of 724 bp in length  
13554 13653: gap of 100 bp  
13654 14359: contig of 706 bp in length  
14360 14459: gap of 100 bp  
14460 15161: contig of 702 bp in length  
15162 15261: gap of 100 bp  
15262 15950: contig of 689 bp in length  
15951 16050: gap of 100 bp  
16051 16749: contig of 699 bp in length  
16750 16849: gap of 100 bp  
16850 17542: contig of 693 bp in length  
17543 17642: gap of 100 bp  
17643 18352: contig of 710 bp in length  
18353 18452: gap of 100 bp  
18453 19168: contig of 716 bp in length  
19169 19268: gap of 100 bp  
19269 19983: contig of 715 bp in length  
19984 20083: gap of 100 bp  
20084 20778: contig of 695 bp in length  
20779 20878: gap of 100 bp  
20879 21577: contig of 699 bp in length  
21578 21677: gap of 100 bp  
21678 22380: contig of 703 bp in length  
22381 22480: gap of 100 bp  
22481 23195: contig of 715 bp in length  
23196 23295: gap of 100 bp  
23296 24011: contig of 716 bp in length  
24012 24111: gap of 100 bp  
24112 24824: contig of 713 bp in length  
24825 24924: gap of 100 bp  
24925 25636: contig of 712 bp in length  
25637 25736: gap of 100 bp  
25737 26448: contig of 712 bp in length  
26449 26548: gap of 100 bp  
26549 27255: contig of 707 bp in length  
27256 27355: gap of 100 bp  
27356 28061: contig of 706 bp in length  
28062 28161: gap of 100 bp  
28162 28866: contig of 705 bp in length  
28867 28966: gap of 100 bp  
28967 29664: contig of 698 bp in length  
29665 29764: gap of 100 bp  
29765 30465: contig of 701 bp in length  
30466 30565: gap of 100 bp  
30566 31275: contig of 710 bp in length  
31276 31375: gap of 100 bp  
31376 32084: contig of 709 bp in length  
32085 32184: gap of 100 bp  
32185 32874: contig of 690 bp in length  
32875 32974: gap of 100 bp  
32975 33679: contig of 705 bp in length  
33680 33779: gap of 100 bp  
33780 34485: contig of 706 bp in length  
34486 34585: gap of 100 bp  
34586 35279: contig of 694 bp in length  
35280 35379: gap of 100 bp  
35380 36062: contig of 683 bp in length  
36063 36162: gap of 100 bp  
36163 36856: contig of 694 bp in length  
36857 36956: gap of 100 bp  
36957 37656: contig of 700 bp in length  
37657 37756: gap of 100 bp  
37757 38460: contig of 704 bp in length  
38461 38560: gap of 100 bp  
38561 39273: contig of 713 bp in length  
39274 39373: gap of 100 bp  
39374 40097: contig of 724 bp in length

40098 40197: gap of 100 bp  
40198 40889: contig of 692 bp in length  
40890 40989: gap of 100 bp  
40990 41692: contig of 703 bp in length  
41693 41792: gap of 100 bp  
41793 42505: contig of 713 bp in length  
42506 42605: gap of 100 bp  
42606 43318: contig of 713 bp in length  
43319 43418: gap of 100 bp  
43419 44124: contig of 706 bp in length  
44125 44224: gap of 100 bp  
44225 44923: contig of 699 bp in length  
44924 45023: gap of 100 bp  
45024 45739: contig of 716 bp in length  
45740 45839: gap of 100 bp  
45840 46534: contig of 695 bp in length  
46535 46634: gap of 100 bp  
46635 47333: contig of 699 bp in length  
47334 47433: gap of 100 bp  
47434 48138: contig of 705 bp in length  
48139 48238: gap of 100 bp  
48239 48937: contig of 699 bp in length  
48938 49037: gap of 100 bp  
49038 49743: contig of 706 bp in length.  
location/Qualifiers  
1. 49743  
/organism="Mus musculus"  
/db\_xref="taxon:10090"  
/clone="RP23-8963"  
/clone\_lib="RP23-8963"  
BASE COUNT 10978 a 10783 c 10620 g 11215 t 6147 others  
ORIGIN  
Query Match 12.18; Score 173; DB 2; Length 49743;  
Best Local Similarity 100.0%; Pred. No. 3,8e-84;  
Matches 173; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 539 GGGCAGTGCAGCGCCAGGCTGACAGCAGCCTCAGTGTGAGAGCAGCAGCAGAGA 598  
|||||  
DB 3661 GGGCAGTGCAGCGCCAGGCTGACAGCAGCCTCAGTGTGAGAGCAGCAGCAGAGA 3602  
|||||  
OY 599 TCACATCTACTGCTGAGTGCAGAGTGCCACCTGCTCTCTGCAAGGTTTGGGCG 658  
|||||  
DB 3601 TCACATCTACTGCTGAGTGCAGAGTGCCACCTGCTCTCTGCAAGGTTTGGGCG 3542  
|||||  
OY 659 CCCACAGGAGCTGTGAGTGGCCCCCTGCCCCACATTTCACAAAGCCCAAG 711  
|||||  
DB 3541 CCCACAGGAGCTGTGAGTGGCCCCCTGCCCCACATTTCACAAAGCCCAAG 3489  
|||||  
RESULT 6  
AC095184  
LOCUS  
DEFINITION  
Rattus norvegicus clone CH230-9D7, \*\*\* SEQUENCING IN PROGRESS \*\*\*  
85 unordered pieces.  
AC095184  
VERSION  
AC095184.3 GI:21722570  
KEYWORDS  
HTG; HTGS; PHASEL.  
SOURCE  
Norway rat.  
ORGANISM  
Rattus norvegicus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.  
1 (bases 1 to 164772)  
REFERENCE  
AUTHORS  
Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,  
Alsbrooks,S.L., Amaratunge,H.C., Are,J.R., Ayale,M., Banks,T.,  
Barbaria,J., Benton,J., Bimage,K., Blankenburg,K., Bonnin,D.,  
Bouck,J., Bowie,S., Brileva,M., Brown,E., Brown,M., Bryant,N.P.,  
Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,  
Carroll,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,  
Chen,G., Chen,R., Chen,Z., Chowdhury,I., Christopoulos,C.,  
Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,  
Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,



Delaney,R.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,  
Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,  
Earhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escoto,M.,  
Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,  
Gabisi,A., Geo,U., Garcia,A., Garner,T., Garza,N., Gill,R.,  
Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,  
Harris,C., Harris,K., Hart,M., Haylak,P., Hawes,A., Hernandez,J.,  
Hernandez,O., Hodgson,A., Hognes,M., Holloway,C., Hollins,B.,  
Homsí,F., Howard,S., Huber,J., Huliyk,S., Hume,J., Jackson,L.E.,  
Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,  
Karlssohn,E., Kelly,S., Khan,U., King,L., Kovach,J., Kovar,C.,  
Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,  
Li,J., Li,Z., Lichtenarge,O., Lieu,C., Liu,J., Liu,W., Loulseged,H.,  
Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,  
Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E.,  
Massey,E., Mawhney,E., Mcleod,M.P., Meador,M., Mei,G., Metzker,M.,  
Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,  
Moser,M., Nickerson,E., Nwokenkwo,S., Ogulu,M., Okwuonu,G.,  
Nguyen,N., Nickerson,E., Nwokenkwo,S., Ogulu,M., Okwuonu,G.,  
Ogunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,  
Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,  
Rives,M., Rojas,A., Rojubokan,I., Rolfe,M., Ruiz,S., Savery,G.,  
Scherer,S., Scott,G., Shen,H., Shooshair,N., Sisson,I.,  
Sodergren,E., Sonalke,T., Sparks,A., Stanley,H., Stone,H.,  
Sutton,A., Svatek,A., Taber,P., Tamerisa,A., Tamerisa,K., Tang,H.,  
Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S.,  
Umanu,K., Vasquez,L., Vera,V., Villalón,D., Vinson,R., Wang,Q.,  
Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S.,  
Williams,G., Williamson,A., Wleczek,R., Wooden,S., Worley,K.,  
Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,  
Weinstock,G. and Gibbs,R.

Direct Submission  
2 (bases 1 to 164772)  
Worley,K.C.

Direct Submission  
Submitted (16-SEP-2001) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA  
3 (bases 1 to 164772)  
Worley,K.C.

Direct Submission  
Submitted (11-JUL-2002) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA  
On Jul 10, 2002 this sequence version replaced gi:17942274.

----- Genome Center  
Center: Baylor College of Medicine  
Center code: BCM  
Web site: <http://www.hgsc.bcm.tmc.edu/>  
Contact: hgsc-help@bcm.tmc.edu  
----- Project Information  
Center project name: GDAW  
Center clone name: CH230-9D7  
----- Summary Statistics  
Sequencing vector: Plasmid  
Chemistry: Dye-terminator Big Dye: 100% of reads  
Assembly program: Phrap: version 0.990329  
Consensus quality: 79000 bases at least Q40  
Consensus quality: 85421 bases at least Q30  
Consensus quality: 90756 bases at least Q20  
-----  
\* NOTE: Estimated insert size may differ from sequence length  
\* (see <http://www.hgsc.bcm.tmc.edu/docs/Genbankdraft.data.html>).  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 85 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.  
1 1297: contig of 1297 bp in length

1298 1397: gap of unknown length  
1398 2679: contig of 1282 bp in length  
2680 2779: gap of unknown length  
2780 3802: contig of 1023 bp in length  
3803 3902: gap of unknown length  
3903 4951: contig of 1049 bp in length  
4952 5051: gap of unknown length  
5052 6299: contig of 1248 bp in length  
6300 6400 7431: gap of unknown length  
7432 7531: gap of unknown length  
7532 8548: contig of 1017 bp in length  
8549 9776: contig of 1128 bp in length  
9777 9876: gap of unknown length  
9877 10877: contig of 1001 bp in length  
10878 10977: gap of unknown length  
10979 12067: contig of 1090 bp in length  
12068 12167: gap of unknown length  
12169 13239: contig of 1072 bp in length  
13240 13339: gap of unknown length  
13340 14779 14778: contig of 1439 bp in length  
14779 14878: gap of unknown length  
14879 16415: contig of 1537 bp in length  
16416 17592: gap of unknown length  
17593 17692: contig of 1077 bp in length  
17693 18970: gap of unknown length  
18971 19070: contig of 1278 bp in length  
19071 20403: gap of unknown length  
20404 20503: contig of 1333 bp in length  
20504 21651: gap of unknown length  
21652 21751: contig of 1148 bp in length  
21752 22887: gap of unknown length  
22887 22987: contig of 1136 bp in length  
22988 22987: gap of unknown length  
22988 24446: contig of 1459 bp in length  
24447 24546: gap of unknown length  
24547 25565: contig of 1019 bp in length  
25566 27036: gap of unknown length  
27037 27136: contig of 1371 bp in length  
27137 28478: gap of unknown length  
28479 28578: contig of 1342 bp in length  
28579 29664: gap of unknown length  
29665 29764: contig of 1086 bp in length  
29765 31068: gap of unknown length  
31069 31168: contig of 1304 bp in length  
31169 32406: gap of unknown length  
32407 32506: contig of 1238 bp in length  
32507 34145: gap of unknown length  
34146 34245: contig of 1639 bp in length  
34246 35811: gap of unknown length  
35812 35911: contig of 1566 bp in length  
35912 37018: gap of unknown length  
37019 37118: contig of 1107 bp in length  
37119 38624: gap of unknown length  
38625 38724: contig of 1506 bp in length  
38725 40409: gap of unknown length  
40410 40509: contig of 1685 bp in length  
40510 42002: gap of unknown length  
42003 42102: contig of 1493 bp in length  
42103 43258: gap of unknown length  
43259 44262: contig of 1056 bp in length  
44263 44362: gap of unknown length  
44363 45668: gap of unknown length  
45669 47130 47129: gap of 1206 bp in length  
47130 47229: contig of 1461 bp in length  
47230 48940: gap of unknown length  
48941 49040: contig of 1711 bp in length  
49041 50571: gap of unknown length  
50572 50671: contig of 1531 bp in length  
50672 50671: gap of unknown length



\* 50672 52207: contig of 1536 bp in length  
\* 52208 52307: gap of unknown length  
\* 52308 52308 52307: gap of unknown length  
\* 53450 53450 53449: gap of 1142 bp in length  
\* 53450 53450 53449: gap of unknown length  
\* 54870 54870 54869: contig of 1320 bp in length  
\* 54870 54870 54869: gap of unknown length  
\* 54870 54870 54869: gap of unknown length  
\* 56027 56027 56026: contig of 1057 bp in length  
\* 56027 56027 56026: gap of unknown length  
\* 56127 56127 56126: contig of 2485 bp in length  
\* 56127 56127 56126: gap of unknown length  
\* 58712 58712 58711: gap of 1331 bp in length  
\* 60043 60043 60042: contig of 1309 bp in length  
\* 60143 60143 60142: gap of unknown length  
\* 61452 61452 61451: contig of 1309 bp in length  
\* 61552 61552 61551: gap of unknown length  
\* 63615 63615 63614: contig of 2063 bp in length  
\* 63715 63715 63714: gap of unknown length  
\* 65157 65157 65156: contig of 1443 bp in length  
\* 65158 65158 65157: gap of unknown length  
\* 65258 65258 65257: gap of unknown length  
\* 66957 66957 66956: gap of 1700 bp in length  
\* 67058 67058 67057: gap of unknown length  
\* 69461 69461 69460: contig of 2404 bp in length  
\* 69462 69462 69461: gap of unknown length  
\* 70953 70953 70952: contig of 1392 bp in length  
\* 70954 70954 70953: gap of unknown length  
\* 71054 71054 71053: gap of unknown length  
\* 72993 72993 72992: contig of 1939 bp in length  
\* 73093 73093 73092: gap of unknown length  
\* 75635 75635 75634: contig of 2542 bp in length  
\* 75734 75734 75733: gap of unknown length  
\* 77845 77845 77844: contig of 2110 bp in length  
\* 77945 77945 77944: gap of unknown length  
\* 77945 77945 77944: gap of unknown length  
\* 80324 80324 80323: gap of 2379 bp in length  
\* 81947 81947 81946: contig of 1524 bp in length

Query Match 4.0%; Score 57; DB 2; Length 164772;  
Best Local Similarity 100.0%; Pred. No. 1.2e-19;  
Matches 57; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 655 GGGCCCAAGAGAGTGTGAGTGGCCCTCGCCACCATTTACAAAGCCAGAG 711  
Db 162273 GGGCCCAAGAGAGTGTGAGTGGCCCTCGCCACCATTTACAAAGCCAGAG 162329

RESULT 7  
LOCUS HSA291714 1329 bp mRNA linear PRI 06-JUL-2001  
DEFINITION Homo sapiens mRNA for RNF30 gene for ring finger protein 30.  
ACCESSION AJ291714  
VERSION AJ291714.1 GI:13160387  
KEYWORDS RNF-3 gene; muscle specific ring finger protein.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 1329)  
Centner,T., Yano,J., Kimura,E., McElhinny,A.S., Pellin,K.,  
Wilt,C.C., Bang,M.L., Trombidas,K., Granzier,H., Gregorio,C.C.,  
Sornichsen,H. and Labeit,S.  
Identification of muscle specific ring finger proteins as potential  
regulators of the titin kinase domain  
J. Mol. Biol. 306 (4), 717-726 (2001)  
JOURNAL MEDLINE 21140140  
PUBMED 11243782  
REFERENCE 2 (bases 1 to 1329)  
Centner,T.  
Direct Submission  
Submitted (12-FEB-2001) Centner T., Structure and Biocomputing,  
EMBL-Heidelberg, Meyerhofstrasse 1, Heidelberg 69112, Germany  
Revised by author 22-FEB-2001  
REMARK FEATURES  
location/Qualifiers  
1..1329  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"

misc-feature 1..3  
/note="putative alternative translation initiation codon"  
/evidence=not\_experimental  
gene 49..1203  
/gene="RNF30"  
CDS 49..1203  
/gene="RNF30"  
/note="alternative"  
/product="ring finger protein 30"  
/protein\_id="CAC32842.1"  
/db\_xref="GI:13160389"

CDS  
/translation="MDNLEKQLICPICLEMFSKPVVILPCQHNLCRCANDVPOASNP  
LWQSRSTVYSGRGRCPSGRHVVLDHGVYGLQNLIVENTIDYIKQSSPRLS  
KAEQHMCEHEEKEKINIVYLSCEVPSLCYFKAGKDEVAFLPTIRKQKODLT  
LPLRLCSGNTTYCSIDLPSSDPPLASQNKIIDSSEQJAMVAGNDVROAVI  
TQMEVFCQRTEDNSRRQKOLLQRFESICVAEEERKGLLOALABQEEKIQRVGLI  
RQYGDHLEASSKIVESAIOSEERPMALYIQAKELINKGAMSKVLAEPREGYES  
MEQFIVRVEHVAEMLRITDFQPGASGEERVAADPGEESAGPEERPDGP"  
join(49..515,642..1203)  
/gene="RNF30"  
/note="short isoform"  
/codon\_start=1  
/product="ring finger protein 30"  
/protein\_id="CAC32841.1"  
/db\_xref="GI:13160388"

BASE COUNT 310 a 375 c 414 g 230 t

Query Match 2.5%; Score 36; DB 9; Length 1329;  
Best Local Similarity 100.0%; Pred. No. 5.1e-08;  
Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1099 CGGCGGAGCCAGGCTATGAGAGCATGAGCAATTC 1134  
Db 1027 CGGCGGAGCCAGGCTATGAGAGCATGAGCAATTC 1062

RESULT 8  
LOCUS AC013413 118847 bp DNA linear PRI 21-FEB-2002  
DEFINITION Homo sapiens BAC clone RP11-538J11 from 2, complete sequence.  
ACCESSION AC013413  
VERSION AC013413.6 GI:18158397  
KEYWORDS HMG.  
SOURCE Homo sapiens.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 118847)  
Sullivan,J.E. and Waterston,R.  
Toward a complete human genome sequence  
Genome Res. 8 (11), 1097-1108 (1998)  
JOURNAL MEDLINE 99063792  
PUBMED 9847074  
REFERENCE 2 (bases 1 to 118847)  
Swearengen,S. and Kozlowski,A.  
The sequence of Homo sapiens BAC clone RP11-538J11  
Unpublished (2001)  
REFERENCE 3 (bases 1 to 118847)  
Waterston,R.H.  
Direct Submission  
Submitted (09-NOV-1999) Genome Sequencing Center, Washington  
University School of Medicine, 4444 Forest Park Parkway, St. Louis,  
MO 63108, USA  
4 (bases 1 to 118847)

AUTHORS Waterston, R. H.  
 TITLE Direct Submission  
 JOURNAL Submitted (16-JAN-2002) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA  
 5 (bases 1 to 118847)  
 WATERSTON, R.  
 DIRECT SUBMISSION  
 Submitted (21-FEB-2002) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA  
 On Jan 10, 2002 this sequence version replaced g1:16418272.  
 -----  
 Center: Washington University Genome Sequencing Center  
 Center code: WUGSC  
 Web site: http://genome.wustl.edu/gsc  
 Contact: saplens@wustl.wustl.edu  
 -----  
 Summary Statistics  
 Center project name: H\_NH0538J11  
 -----

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:  
 Mapping information for this clone was provided by Dr. John D. McPherson, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see http://genome.wustl.edu/gsc

SOURCE INFORMATION:  
 The RPCR-11 human BAC library was made from the blood of one male donor, as described by Osceogawa, K., Woon, P.Y., Zhao, B., Frengen, E., Tatenou, M., Catalanese, J.J. and de Jong, P.J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. Genomics 51:1-8. The clone may be obtained either from Research Genetics, Inc. (http://www.resgen.com) or Pieter de Jong and coworkers at http://www.chori.org  
 VECTOR: pBACE3.6

NEIGHBORING SEQUENCE INFORMATION:  
 The clone sequenced to the left is RP11-195B17, 2000 bp overlap. Actual start of this clone is at base position 80914 of RP11-195B17; actual end is at base position 118847 of RP11-538J11.  
 RP11-538J11 contains a transposon in the growth of the clone which has been omitted from the finished sequence. Data from AC025662 was used to finish this clone, AC013413.

#### FEATURES

source  
 1. 118847  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /chromosome="2"  
 /map="2"  
 /clone="RP11-538J11"  
 /clone\_id="RPCR-11"  
 1. 1159  
 /note="match to EST BE789511 (NID:g10210709)"  
 1. 1151  
 /note="match to EST A1978781 (NID:g5803811) wr59f04.x1"  
 133. 154  
 /note="match to EST BG256909 (NID:g12766725)"  
 393. 626  
 /note="similar to Mus musculus EST BB659249

misc\_feature (NID:g16493072)"  
 404. 626  
 /note="similar to Mus musculus EST B1661519 (NID:g1557575)"  
 691. 700  
 /note="match to EST A1905917 (NID:g6496304)"  
 870. 1009  
 /note="similar to Mus musculus EST BB659249 (NID:g16493072)"  
 870. 1009  
 /note="similar to Mus musculus EST B1661519 (NID:g1557575)"  
 1991. 2327  
 /note="match to EST AA758920 (NID:g2806783) ah77b09.sl"  
 2336. 2554  
 /rpl\_family="Alu"  
 2551. 2600  
 /rpl\_family="T-rich"  
 2576. 2880  
 /rpl\_family="Alu"  
 2971. 3156  
 /rpl\_family="L2"  
 3359. 3667  
 /rpl\_family="Alu"  
 3669. 3979  
 /rpl\_family="Alu"  
 3951. 3979  
 /rpl\_family="AT-rich"  
 4211. 4340  
 /note="similar to Mus musculus EST BB659249 (NID:g16493072)"  
 4211. 4340  
 /note="similar to Mus musculus EST B1661519 (NID:g1557575)"  
 4529. 5049  
 /rpl\_family="L1"  
 4747. 4773  
 /rpl\_family="(T)n"  
 5187. 5329  
 /note="similar to Mus musculus EST B1661519 (NID:g1557575)"  
 5187. 5309  
 /note="similar to Mus musculus EST BB659249 (NID:g16493072)"  
 5512. 5651  
 /note="similar to Mus musculus EST B1661519 (NID:g1557575)"  
 5528. 5657  
 /note="similar to Homo sapiens EST BF761009 (NID:g12108987)"  
 5552. 5653  
 /note="match to EST BE242145 (NID:g9093794)"  
 5552. 5653  
 /note="match to EST BM152742 (NID:g17177267)"  
 5553. 5653  
 /note="similar to Mus musculus EST BF532585 (NID:g11619948)"  
 5859. 6030  
 /note="match to EST BM152742 (NID:g17177267)"  
 5859. 6030  
 /note="similar to Mus musculus EST BF532585 (NID:g11619948)"  
 5859. 5931  
 /note="match to EST BE242145 (NID:g9093794)"  
 5867. 5912  
 /note="similar to Homo sapiens EST BF761009 (NID:g12108987)"  
 5942. 6030  
 /note="match to EST AT244159 (NID:g3839556) qj80e03.x1"  
 6145. 6450  
 /rpl\_family="Alu"  
 6428. 6450  
 /note="match to EST BE789511 (NID:g10210709)"

misc-feature 6556..6741  
 /note="match to EST AI244159 (NID:q3839556) qj80e03.x1"  
 misc-feature 6556..6741  
 /note="similar to Mus musculus EST BF532585  
 (NID:g11619948)"  
 misc-feature 6556..6692  
 /note="match to EST BM152742 (NID:g17177267)"  
 misc-feature 6910..7022  
 /note="match to EST AI244159 (NID:q3839556) qj80e03.x1"  
 misc-feature 6910..7022  
 /note="similar to Mus musculus EST BF532585  
 (NID:g11619948)"  
 repeat-region 7252..7332  
 /rpt\_family="MIR"  
 misc-feature 7338..7483  
 /note="similar to Mus musculus EST BF532585  
 (NID:g11619948)"  
 misc-feature 7338..7420  
 /note="match to EST AI244159 (NID:q3839556) qj80e03.x1"  
 misc-feature 7738..7817  
 /note="similar to Mus musculus EST BF532585  
 (NID:g11619948)"  
 repeat-region 8289..8603  
 /rpt\_family="Alu"  
 repeat-region 8292..8316  
 /rpt\_family="(?)n"  
 misc-feature 8866..8919  
 /note="similar to Mus musculus EST B1647507"

Query Match 2.5%; Score 36; DB 9; Length 118847;  
 Best Local Similarity 100.0%; Pred. No. 5.7e-08;  
 Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1099 CGCCGAGCAGGCTATGAGACATGACCAATTC 1134  
 ||||||||||||||||||||||||||||||||||||  
 DB 89240 CGCCGAGCAGGCTATGAGACATGACCAATTC 89275

RESULT 9  
 LOCUS AX060632 1500 bp DNA linear PAT 22-JAN-2001  
 DEFINITION Sequence 54 from Patent WO0078954.  
 ACCESSION AX060632  
 VERSION AX060632.1 GI:12406053  
 KEYWORDS  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Euthera; Primates; Catarrhini; Homiidae; Homo.

REFERENCE 1 (bases 1 to 1500)  
 AUTHORS Lal,P., Yue,H., Tang,Y.T., Baughn,M.R., Azimzai,Y. and Tran,B.  
 TITLE Human transcriptional regulator proteins  
 JOURNAL Patent: WO 0078954-A 54 28-DEC-2000;  
 Inocyte Genomics, Inc. (US)  
 FEATURES  
 Location/Qualifiers  
 1..1500  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"

BASE COUNT 401 a 378 c 439 g 282 t  
 ORIGIN

Query Match 2.4%; Score 35; DB 6; Length 1500;  
 Best Local Similarity 100.0%; Pred. No. 1.8e-07;  
 Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 685 CTGCCACCATTTACAAAGCCAGAGAGTACT 719  
 ||||||||||||||||||||||||||||||||||||  
 DB 777 CTGCCACCATTTACAAAGCCAGAGAGTACT 811

RESULT 10  
 LOCUS AC025662/c 123280 bp. DNA linear HTG 13-APR-2000

DEFINITION Homo sapiens clone RP11-352122, LOW-PASS SEQUENCE SAMPLING.  
 AC025662  
 VERSION 2 GI:7547186  
 HTG: HTGS PHASE0.  
 KEYWORDS  
 SOURCE Homo sapiens.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Euthera; Primates; Catarrhini; Homiidae; Homo.

REFERENCE 1 (bases 1 to 123280)  
 AUTHORS Birren,B., Linton,L., Nusbaum,C. and Lander,E.  
 TITLE  
 JOURNAL  
 REFERENCE  
 AUTHORS

2 (bases 1 to 123280)  
 Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,  
 Anderson,S., Baldwin,J., Barna,N., Bastien,V., Bede,F.,  
 Boguslavsky,L., Bouckgalter,B., Brown,A., Burkett,G.,  
 Campoliano,A., Castle,A., Chappel,Y., Colangelo,M., Collins,S.,  
 Collamore,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S.,  
 Dodge,S., Domino,M., Doyle,M., Ferreira,P., Fitzhugh,W., Gage,D.,  
 Galagan,J., Gardyna,S., Glnde,S., Goyette,M., Graham,L.,  
 Grand-pierre,N., Grant,G., Hagos,B., Heatford,A., Horton,L.,  
 Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,  
 Klein,J., Labocque,K., Lamazares,R., Landers,T., Lechoczy,J.,  
 Levine,R., Lieu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N.,  
 McCarthy,M., McEwan,P., McCurt,A., McKernan,K., McSheeters,R.,  
 Merdum,J., Meneus,L., Mihova,T., Miranda,C., Mienga,V., Morrow,J.,  
 Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,  
 O'Neil,D., Oliver,T.M., Oliver,J., Peterson,C., Plerre,N.,  
 Pisanil,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,  
 Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,  
 Stange-Thoman,N., Stojanovic,N., Subramanian,A., Talamas,J.,  
 Tesfaye,S., Theodore,J., Tirrell,A., Travers,M., Trigilio,J.,  
 Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,  
 Young,G., Zainoun,J., Zimmer,A. and Zody,M.  
 Direct Submission  
 Submitted (12-MAR-2000) Whitehead Institute/MIT Center for Genome  
 Research, 320 Charles Street, Cambridge, MA 02141, USA  
 On Apr 13, 2000 this sequence version replaced gi:7230272.  
 All repeats were identified using RepeatMasker:  
 Smit, A.F.A. & Green, P. (1996-1997)  
 http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center  
 Center: Whitehead Institute/ MIT Center for Genome Research  
 Center code: WIBR  
 Web site: http://www-seq.wi.mit.edu  
 Contact: sequence.submissions@genome.wi.mit.edu  
 ----- Project Information  
 Center project name: 352\_1-22  
 Center clone name: 352\_1-22

\* NOTE: This record contains 146 individual  
 \* sequencing reads that have not been assembled into  
 \* contigs. Runs of N are used to separate the reads  
 \* and the order in which they appear is completely  
 \* arbitrary. Low-pass sequence sampling is useful for  
 \* identifying clones that may be gene-rich and allows  
 \* overlap relationships among clones to be deduced.  
 \* However, it should not be assumed that this clone  
 \* will be sequenced to completion. In the event that  
 \* the record is updated, the accession number will  
 \* be preserved.  
 \* 1  
 \* 749 848: gap of 100 bp  
 \* 849 1619: contig of 771 bp in length  
 \* 1620 1719: gap of 100 bp  
 \* 1720 2490: contig of 771 bp in length  
 \* 2491 2590: gap of 100 bp  
 \* 2591 3378: contig of 788 bp in length  
 \* 3379 3478: gap of 100 bp  
 \* 3479 4239: contig of 761 bp in length  
 \* 4240 4339: gap of 100 bp  
 \* 4340 5113: contig of 774 bp in length  
 \* 5114 5213: gap of 100 bp

\* 5214 5994: contig of 781 bp in length  
\* 5995 6094: gap of 100 bp  
\* 6095 6873: contig of 779 bp in length  
\* 6874 6973: gap of 100 bp  
\* 6974 7732: contig of 759 bp in length  
\* 7733 7832: gap of 100 bp  
\* 7833 8604: contig of 772 bp in length  
\* 8605 8704: gap of 100 bp  
\* 8705 9478: contig of 774 bp in length  
\* 9479 9578: gap of 100 bp  
\* 9579 10356: contig of 778 bp in length  
\* 10357 10456: gap of 100 bp  
\* 10457 11243: contig of 787 bp in length  
\* 11244 11343: gap of 100 bp  
\* 11344 12125: contig of 782 bp in length  
\* 12126 12225: gap of 100 bp  
\* 12226 13028: contig of 803 bp in length  
\* 13029 13128: gap of 100 bp  
\* 13129 13891: contig of 763 bp in length  
\* 13892 13991: gap of 100 bp  
\* 13992 14763: contig of 772 bp in length  
\* 14764 14863: gap of 100 bp  
\* 14864 15625: contig of 762 bp in length  
\* 15626 15725: gap of 100 bp  
\* 15726 16498: contig of 773 bp in length  
\* 16499 16598: gap of 100 bp  
\* 16599 17376: contig of 778 bp in length  
\* 17377 17476: gap of 100 bp  
\* 17477 18246: contig of 770 bp in length  
\* 18247 18346: gap of 100 bp  
\* 18347 19137: contig of 791 bp in length  
\* 19138 19237: gap of 100 bp  
\* 19238 20026: contig of 789 bp in length  
\* 20027 20126: gap of 100 bp  
\* 20127 20909: contig of 783 bp in length  
\* 20910 21009: gap of 100 bp  
\* 21010 21787: contig of 778 bp in length  
\* 21788 21887: gap of 100 bp  
\* 21888 22671: contig of 784 bp in length  
\* 22672 22771: gap of 100 bp  
\* 22772 23556: contig of 785 bp in length  
\* 23557 23656: gap of 100 bp  
\* 23657 24446: contig of 790 bp in length  
\* 24447 24546: gap of 100 bp  
\* 24547 25324: contig of 778 bp in length  
\* 25325 25424: gap of 100 bp  
\* 25425 26211: contig of 787 bp in length  
\* 26212 26311: gap of 100 bp  
\* 26312 27074: contig of 763 bp in length  
\* 27075 27174: gap of 100 bp  
\* 27175 27960: contig of 786 bp in length  
\* 27961 28060: gap of 100 bp  
\* 28061 28825: contig of 765 bp in length  
\* 28826 28925: gap of 100 bp  
\* 28926 29722: contig of 797 bp in length  
\* 29723 29822: gap of 100 bp  
\* 29823 30598: contig of 776 bp in length  
\* 30599 30698: gap of 100 bp  
\* 30699 31472: contig of 774 bp in length  
\* 31473 31572: gap of 100 bp  
\* 31573 32358: contig of 786 bp in length  
\* 32359 32458: gap of 100 bp  
\* 32459 33227: contig of 769 bp in length  
\* 33228 33327: gap of 100 bp  
\* 33328 34098: contig of 771 bp in length  
\* 34099 34198: gap of 100 bp  
\* 34199 34978: contig of 780 bp in length  
\* 34979 35078: gap of 100 bp  
\* 35079 35867: contig of 789 bp in length  
\* 35868 35967: gap of 100 bp  
\* 35968 36782: contig of 815 bp in length  
\* 36783 36882: gap of 100 bp  
\* 36883 37630: contig of 748 bp in length

\* 37631 37730: gap of 100 bp  
\* 37731 38499: contig of 769 bp in length  
\* 38500 38599: gap of 100 bp  
\* 38600 39388: contig of 789 bp in length  
\* 39389 39488: gap of 100 bp  
\* 39489 40255: contig of 767 bp in length  
\* 40256 40355: gap of 100 bp  
\* 40356 41129: contig of 774 bp in length  
\* 41130 41229: gap of 100 bp  
\* 41230 41998: contig of 769 bp in length  
\* 41999 42098: gap of 100 bp  
\* 42099 42868: contig of 770 bp in length  
\* 42869 42968: gap of 100 bp  
\* 42969 43746: contig of 778 bp in length  
\* 43747 43846: gap of 100 bp  
\* 43847 44632: contig of 786 bp in length  
\* 44633 44732: gap of 100 bp  
\* 44733 45522: contig of 790 bp in length  
\* 45523 45622: gap of 100 bp  
\* 45623 46403: contig of 781 bp in length  
\* 46404 46503: gap of 100 bp  
\* 46504 47312: contig of 809 bp in length  
\* 47313 47412: gap of 100 bp  
\* 47413 48189: contig of 777 bp in length  
\* 48190 48289: gap of 100 bp  
\* 48290 49052: contig of 763 bp in length  
\* 49053 49152: gap of 100 bp  
\* 49153 49929: contig of 777 bp in length  
\* 49930 50029: gap of 100 bp  
\* 50030 50795: contig of 766 bp in length  
\* 50796 50895: gap of 100 bp  
\* 50896 51681: contig of 786 bp in length  
\* 51682 51781: gap of 100 bp  
\* 51782 52563: contig of 782 bp in length  
\* 52564 52663: gap of 100 bp  
\* 52664 53443: contig of 780 bp in length  
\* 53444 53543: gap of 100 bp  
\* 53544 54331: contig of 788 bp in length  
\* 54332 54431: gap of 100 bp  
\* 54432 55216: contig of 785 bp in length  
\* 55217 55316: gap of 100 bp  
\* 55317 56115: contig of 799 bp in length  
\* 56116 56215: gap of 100 bp  
\* 56216 57005: contig of 790 bp in length  
\* 57006 57105: gap of 100 bp  
\* 57106 57874: contig of 769 bp in length  
\* 57875 57974: gap of 100 bp  
\* 57975 58750: contig of 776 bp in length  
\* 58751 58850: gap of 100 bp  
\* 58851 59618: contig of 768 bp in length  
\* 59619 59718: gap of 100 bp  
\* 59719 60490: contig of 772 bp in length  
\* 60491 60590: gap of 100 bp

Query Match 2.0%; Score 29; DB 2; Length 123280;  
Best Local Similarity 100.0%; Pred. No. 0.00044;  
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 622 GAGTGGCCACCTGCTCTCTGCAAGT 650  
DB 32873 GAGTGGCCACCTGCTCTCTGCAAGT 32845  
RESULT 11  
AX418852 1597 bp DNA linear PAT 18-JUN-2002  
LOCUS AX418852  
DEFINITION Sequence 5 from Patent WO0206318.  
ACCESSION AX418852  
VERSION AX418852.1 GI:21523716  
KEYWORDS house mouse.  
SOURCE Mus musculus  
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE 1 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 AUTHORS Olson, E.N. and Spencer, J.A.  
 TITLE Methods and compositions for stabilizing microtubules and intermediate filaments in striated muscle cells  
 JOURNAL Patent: WO 0206318-A 5 24-JAN-2002;  
 BOARD OF REGENTS THE UNIVERSITY OF TEXAS SYSTEM (US)  
 FEATURES Location/Qualifiers  
 source 1. 1597  
 /organism="Mus musculus"  
 /db\_xref="taxon:10090"  
 299..1330  
 /note="unnamed protein product"  
 /protein\_id="CAD35457.1"  
 /cdon\_start=1  
 /db\_xref="GI:21523717"  
 /translation="MDYKSSLIIPDGNMENEKOLICPICLEMFKPVILLPCQNLIC  
 RKANDIFFOANPYWTRNGSVSMGSRFRPSCREHYIMDRHGYIGLQRLNLVNI  
 DIYKQECSSRPLOKSHPMCKEHEDEKINITYCLCEVPTCSLCVYGAHQAEVAPLO  
 SITQOGKTELSTNCISMLVAGNDRVQITISQLEDSRVKRENSHOKBELSOKFDPYLA  
 ILDEKKSSELLQRIQEOEENKIGFTEALIIQYREOLEKSTKIVETALISQLEDEGATFL  
 SSKOLIKSIYEASKGCGQLGTEGCFEMDMYFTLDLEHIAELRALRIDGTGKGCVCVC  
 LTFEERSSS"  
 BASE COUNT 405 a 404 c 438 g 350 t  
 ORIGIN  
 Query Match 2.0%; Score 28; DB 6; Length 1597;  
 Best Local Similarity 100.0%; Pred. No. 0.0014;  
 Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 505 AACATCATGTGACATCTACACAGCAGACT 532  
 ||||||||||||||||||||||||||||  
 Db 596 AACATCATGTGACATCTACACAGCAGACT 623  
 RESULT 12  
 LOCUS AL627314 196461 bp DNA linear ROD 09-JUL-2002  
 DEFINITION Mouse DNA sequence from clone RP23-354H24 on chromosome 4, complete  
 sequence.  
 ACCESSION AL627314  
 VERSION AL627314.6 GI:212727344  
 KEYWORDS HTG.  
 SOURCE house mouse.  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 196461)  
 REFERENCE Direct Submission  
 AUTHORS Submitted (09-JUL-2002) Wellcome Trust Sanger Institute, Hinxton,  
 TITLE Cambridge, UK. E-mail enquiries:  
 JOURNAL humquery@sanger.ac.uk  
 COMMENT On Jul 10, 2002 this sequence version replaced g1:18152385.  
 During sequence assembly data is compared from overlapping clones.  
 Where differences are found these are annotated as variations  
 together with a note of the overlapping clone name. Note that the  
 variation annotation may not be found in the sequence submission  
 corresponding to the overlapping clone, as we submit sequences with  
 only a small overlap as described above.  
 This sequence was finished as follows unless otherwise noted: all  
 regions were either double-stranded or sequenced with an alternate  
 chemistry or covered by high quality data (i.e., phred quality >=  
 30); an attempt was made to resolve all sequencing problems, such  
 as compressions and repeats; all regions were covered by at least  
 one plasmid subclone or more than one M13 subclone; and the  
 assembly was confirmed by restriction digest. The following  
 abbreviations are used to associate primary accession numbers given  
 in the feature table with their source databases: Em: EMBL; Sw: SWISSPROT; Tr: TREMBL; Wp: WORMPEP; Information on the WORMPEP  
 database can be found at  
 http://www.sanger.ac.uk/Projects/C\_elegans/wormpep  
 from the RPI-23 mouse PAC Library

constructed by the group of Pieter de Jong.  
 For further details see <http://www.chori.org/bacpac/home.htm>  
 VECTOR: PBACE3.6  
 ----- Genome Center  
 Center: Wellcome Trust Sanger Institute  
 Center code: SC  
 Web site: <http://www.sanger.ac.uk>  
 Contact: humquery@sanger.ac.uk  
 -----  
 FEATURES  
 source Location/Qualifiers  
 1. 196461  
 /organism="Mus musculus"  
 /db\_xref="taxon:10090"  
 /chromosome="4"  
 /clone="RP23-354H24"  
 /clone\_lib="RPCR-23"  
 BASE COUNT 51927 a 47767 c 46054 g 50713 t  
 ORIGIN  
 Query Match 2.0%; Score 28; DB 10; Length 196461;  
 Best Local Similarity 100.0%; Pred. No. 0.0016;  
 Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 505 AACATCATGTGACATCTACACAGCAGACT 532  
 ||||||||||||||||||||||||||||  
 Db 188387 AACATCATGTGACATCTACACAGCAGACT 188414  
 RESULT 13  
 LOCUS AB047601 1746 bp mRNA linear PRI 11-OCT-2001  
 DEFINITION Macaca fascicularis brain cDNA, clone:Onpa-10466.  
 ACCESSION AB047601  
 VERSION AB047601.1 GI:9929936  
 KEYWORDS fis (full insert sequence).  
 SOURCE Macaca fascicularis adult male brain parietal lobe cDNA to mRNA,  
 clone\_lib:macaque brain cDNA library Onpa clone:Onpa-10466.  
 ORGANISM Macaca fascicularis  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;  
 Cercopithecinae; Macaca.  
 1  
 REFERENCE Osada, N., Hida, M., Kusuda, J., Tanuma, R., Iseki, K., Hirata, M.,  
 Suto, Y., Hirai, M., Terao, K., Suzuki, Y., Sugano, S. and Hashimoto, K.  
 TITLE Assignment of 118 novel cDNAs of cynomolgus monkey brain to human  
 chromosomes  
 JOURNAL Gene 275 (1), 31-37 (2001)  
 MEDLINE 21458551  
 REFERENCE 2 (bases 1 to 1746)  
 AUTHORS Hashimoto, K., Osada, N., Hida, M., Kusuda, J. and Sugano, S.  
 TITLE Direct Submission  
 JOURNAL Submitted (23-AUG-2000) Katsuyuki Hashimoto, National Institute of  
 Infectious Diseases, Division of Genetic Resources; 23-1, Toyama  
 1-chome, Shinjuku-ku, Tokyo 162-8640, Japan  
 (E-mail: khashim@nih.go.jp, URL: <http://www.nih.go.jp/yoken/genbank/>,  
 Tel: 81-3-5285-1111 (ex. 2120), Fax: 81-3-5285-1181)  
 lab host: TOP10  
 Vector: pME18S-Fl3 (ACC. NO. AB009864)  
 R. Site1: DraIII (CACTGTGTG)  
 R. Site2: DraIII (CACCATGTG)  
 DESCRIPTION: 1st strand cDNA was primed with an oligo(dT) primer  
 [ATGCGCCCTTTTCTTTTCTTTTCTTTT]; double-stranded cDNA was synthesized  
 using specific 5' and 3' primers and amplified by PCR. The PCR  
 product was digested with SfiI and size selection was performed to  
 exclude fragments <1.5Kb. The SfiI-digested PCR product was cloned  
 into distinct DraIII sites of pME18S-Fl3. XhoI sites just outside  
 the DraIII sites can be used to isolate the cDNA insert. Libraries  
 were constructed by Sugano et al. (University of Tokyo, Institute of  
 Medical Science). Custom primer used for sequencing  
 (5' end primer [CTTCTGCTCTAAAGCTGCG];  
 3' end primer [CGACCTGACCTCGAGCACA]).  
 Location/Qualifiers  
 1. 1746  
 source

```

/organism="Macaca fascicularis"
/db_xref="taxon:9541"
/clone="Onpa-10466"
/sex="male"
/tissue_type="brain parietal lobe"
/clone_lib="macaque brain cDNA library Onpa"
/dev_stage="adult"
567. .1244

```

BASE COUNT ORIGIN	523 a	396 c	387 g	440 t
----------------------	-------	-------	-------	-------

Query Match	1.9%;	Score 27;	DB 9;	Length 1746;
Best Local Similarity	100.0%;	Pred. No. 0.0052;		
Matches	27;	Conservative	0;	Mismatches
			0;	Indels
				0;
				Gaps
				0;

<b>QY</b>	<b>508</b>	<b>ATCATGTGACATCTACAGCAGAGTGCC</b>	<b>534</b>
<b>Db</b>	<b>876</b>	<b>ATCATGTGACATCTACAGCAGAGTGCC</b>	<b>902</b>

RESULT	14
LOCUS	BC007750
DEFINITION	Homo sapiens, clone MGC:12836 IMAGE:4110789, mRNA, complete cds.
ACCESSION	BC007750
VERSION	BC007750.1 GI:14043531
KEYWORDS	MGC.
SOURCE	Homo sapiens.
ORGANISM	Homo sapiens.

REFERENCE	1 (pages 1 to 1750)
AUTHORS	Strausberg, R.
TITLE	Direct Submission
JOURNAL	Submitted (11-MAY-2001) National Institutes of Health, Mammalian

REMARK  
NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
Contact: MGC help desk  
COMMENT

CDNA Library Preparation: Rubin Laboratory  
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: National Institutes of Health Intramural  
 Sequencing Center (NISC),  
 Gaithersburg, Maryland;  
 Web site: [http://www.nisc.nih.gov/nisc\\_mgc/hgtrl.nih.gov](http://www.nisc.nih.gov/nisc_mgc/hgtrl.nih.gov)  
 Contact: [nisc\\_mgc/hgtrl.nih.gov](mailto:nisc_mgc/hgtrl.nih.gov)  
 Sherchenko, Y., Wechterby, R.D., Beckström-Sternberg, S.M.,  
 Benjamini, B., Blakesley, R.W., Bouffard, G.G., Brinkley, C.,  
 Dietrich, N.L., Guu, X., Gupta, J., Ho, S.-L., Karlins, E., Legaspi, R.,  
 Lim, M., Maduro, Q.L., Mastello, C., Mastiani, S.D., McCloskey, J.C.,  
 McDowell, J., Pearson, R., Snyder, B., Stantitrop, S., Thomas, P.J.,  
 Thompson, E.E., Touchman, J.W., Tsurgou, C., Vogt, L.L., Walker, M.A.,  
 Zhang, L.-H. and Green, E.D.

## FEATURES

Source

```

source
1..1750
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="MGC:12836 IMAGE:4110783"
/tissue_type="Muscle, rhabdomyosarcoma"
/clone_id="NH_MGC_17"
/lab_host="DH10B-R"
/note="Vector: pOTB7"
190..1500

```

BASE COUNT ORIGIN	556 a	389 c	433 g	372 t
----------------------	-------	-------	-------	-------

Query Match	1.9%;	Score 27;	DB 9;	Length 1750;
Best Local Similarity	100.0%;	Pred. No. 0.0052;		
Matches	27;	Conservative	0;	Mismatches
			0;	Indels
				0;
				Gaps
				0;

Oy	ATCATGTGACATCTACAGCAGAGTCC	534
Db	ATCATTGACATCTACAGCAGAGTCC	477

RESULT	15
accession	HSA243488
LOCUS	HSA243488
DEFINITION	Homo sapiens titin zinc-finger anchoring protein, 50kDa isoform.
ACCESSION	AJ243488
VERSION	AJ243488.1
KEYWORDS	GI:14588845
	alternative splicing; signal transduction; titin zinc-finger
SOURCE	anchoring protein; titian.
	human.

ORGANISM	REFERENCE
Homo sapiens	1 (bases 1 to 1810)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.	

AUTHORS	Kelly, R., Neubauer, G. and Gautel, M.
TITLE	A novel RING finger protein associated with titin kinase
JOURNAL	Unpublished
REFERENCE	2 (bases 1 to 1810)

TITLE	Direct Submission
JOURNAL	Submitted (29-JUN-1999) Gautel M.S., Physikalische Biochemie,

FEATURES

SOURCE

CDS

```
/db_xref="taxon:9606"
/cell_type="cardiomyocyte"
/tissue_type="cardiac muscle"
227..1585
/function="putative role in signal transduction"
/note="50KDa isoform"
/codon_start=1
/evidence=experimental
/product="titin zinc finger anchoring protein"
/protein_id="Cac43019.1"
/db_xref="GI:14588846"
/translation="MSASLWYKSFSEKQQTMDNLKOLICPICLEMTTKRPVILPCOH
NLCKCASDIFQASNPYLPTIRGGTTMAGSGFRCPSRHEVYLDPRHVGKORLLVGG
NIIDIKYQESRPERKSKDOPMCEHEERINIVCLNEVEPLCSIKVEGAKNDQVAP
LTHVFOKSKSLSDGILILVGSNDRVGVISQLEDPTKTIIECCRKQKQELCEKFYIL
```

YGLIEKRNEMTOVITRTOEKLHVRALLIKKYSIDHLENVSKLVESGLOFMEDEPMAY  
FLONAKTLKLIKISEASKAFOMEKIEHGYENNNHFTVINRBEKIIREIDFYREDEDE  
EEEGEGEKEEVEGEAVEVEENVOPEPEDEPEKASELSOVELOAPGALP  
VSSPEPPALPPADAPVTOIGFAPPLQGAAPASGADSEFARHIFSWNLNL  
NE"

BASE COUNT  
ORIGIN

557 a 420 c 448 g 385 t

Query Match 1.9%; Score 27; DB 9; Length 1810;  
Best Local Similarity 100.0%; Pred. No. 0.0052;  
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 508 ATCATTCATCTACAGCAGAGTCC 534  
|||||  
Db 536 ATCATTCATCTACAGCAGAGTCC 562

RESULT 16  
LOCUS

AK091728 1925 bp mRNA linear PRI 15-JUL-2002  
Homo sapiens CDNA FLJ34409 fis, clone HEART2001931, moderately  
similar to Mus musculus RING-finger protein MURF mRNA.

ACCESSION AK091728  
VERSION AK091728.1 GI:21750167  
KEYWORDS oligo capping; fis (full insert sequence).  
SOURCE Homo sapiens heart CDNA to mRNA, clone\_lib:HEART2  
clone:HEART2001931.

ORGANISM

Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE  
AUTHORS

1 Nishi,T., Ota,T., Nakagawa,S., Senoh,A., Mizuguchi,H., Inagaki,H.,  
Sugiyama,T., Irie,R., Otsuki,T., Sato,H., Wakamatsu,A., Ishii,S.,  
Yamamoto,J., Isono,Y., Kawai-Hio,Y., Saito,K., Nishikawa,T.,  
Kimura,K., Yamashita,H., Matsuo,K., Nakamura,Y., Sekine,M.,  
Kikuchi,H., Kanda,K., Magatsuwa,M., Murakawa,K., Kanehori,K.,  
Takahashi-Fujii,A., Oshima,A., Sugiyama,A., Kawakami,B., Suzuki,Y.,  
Sugano,S., Nagahara,K., Masuno,Y., Nagai,K. and Isogai,T.  
NDO human CDNA sequencing project

Unpublished  
2 (bases 1 to 1925)

Isogai,T. and Yamamoto,J.  
Direct Submission

Submitted (04-JUL-2002) Takao Isogai, FLJ Project(HRI Team); 2-6-7  
Kazusa-Kamatairi, Kisarazu, Chiba 292-0812, Japan  
(E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)

COMMENT

NDO human CDNA sequencing project supported by Ministry of  
Economy, Trade and Industry of Japan; CDNA full insert sequencing:  
Research Association for Biotechnology (RAB); CDNA library  
construction: Helix Research Institute (HRI) (supported by Japan  
Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB,  
HRI, and Biotechnology Center, National Institute of Technology and  
Evaluation; clone selection for full insert sequencing: HRI and  
RAB; annotation: HRI and RAB.

FEATURES  
source

1. 1925  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="HEART2001931"  
/tissue\_type="heart"  
/clone\_lib="HEART2"  
/note="cloning vector: PME18SFL3"

BASE COUNT 579 a 441 c 498 g 407 t  
ORIGIN

Query Match 1.9%; Score 27; DB 9; Length 1925;  
Best Local Similarity 100.0%; Pred. No. 0.0052;  
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 508 ATCATTCATCTACAGCAGAGTCC 534  
|||||  
Db 356 ATCATTCATCTACAGCAGAGTCC 382

RESULT 17  
LOCUS

HS243489 2098 bp mRNA linear PRI 29-JUN-2001  
Homo sapiens titin zinc-finger anchoring protein, 60kDa isoform.

ACCESSION  
VERSION

AJ243489.1 GI:14588847

KEYWORDS

alternative splicing; signal transduction; titin zinc-finger  
anchoring protein; titian.

SOURCE

human.

ORGANISM

Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE  
AUTHORS

1 (bases 1 to 2098)  
Kelly,R., Neubauer,G. and Gautel,M.

TITLE

A novel RING finger protein associated with titin kinase

JOURNAL

Unpublished  
2 (bases 1 to 2098)

REFERENCE  
AUTHORS

Gautel,M.S.  
Direct Submission

TITLE

Submitted (29-JUN-1999) Gautel M.S., Physiologische Biochemie,  
Max-Planck-Institut fuer molekulare Physiologie, Otto-Hahn-Strasse  
11, Dortmund, 44227, GERMANY

JOURNAL

Location/Qualifiers  
1. 2098

FEATURES  
source

/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/cell\_type="cardiomyocyte"  
/tissue\_type="cardiac muscle"  
227. 1873  
/function="putative role in signal transduction"  
/note="60kDa isoform"

CDS

/evidence="experimental"  
/product="titin zinc-finger anchoring protein"  
/protein\_id="CA433020.1"  
/db\_xref="GI:14588848"  
/translation="MSASLVNYSFSPKRODMNLEKOLICPLEMFTKPVITLPCOH  
NICKRKAADIFQASNYLPRGCTYMASGGRFRCPSRHVYVLDRAHVGKQVITLPCOH  
NITDIYKQESTPREKSDQPCHEHEERINITLNCVFPICSLCKYFGAHKDCOVAP  
LHVFOROKSELSVDGIALIVGSNDRVGVLSQEDTCTIECCRKQKQELCEKEDVL  
YGLIEKRNEMTOVITRTOEKLHVRALLIKKYSIDHLENVSKLVESGLOFMEDEPMAY  
FLONAKTLKLIKISEASKAFOMEKIEHGYENNNHFTVINRBEKIIREIDFYREDEDE  
EEEGEGEKEEVEGEAVEVEENVOPEPEDEPEKASELSOVELOAPGALP  
VSSPEPPALPPADAPVTOIGFAPPLQGAAPASGADSEFARHIFSWNLNLNE"

BASE COUNT 632 a 497 c 534 g 435 t

ORIGIN

Query Match 1.9%; Score 27; DB 9; Length 2098;  
Best Local Similarity 100.0%; Pred. No. 0.0052;  
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 508 ATCATTCATCTACAGCAGAGTCC 534  
|||||  
Db 536 ATCATTCATCTACAGCAGAGTCC 562

RESULT 18  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

HS243489 2202 bp mRNA linear PRI 06-JUL-2001  
Homo sapiens mRNA for RNF29 gene for ring finger protein 29.  
AJ291712  
AJ291712.1 GI:13160384  
ring finger protein 29; RNF29 gene.  
human.

REFERENCE  
AUTHORS

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 2202)  
Centner,T., Yano,D., Kimura,E., McElhinny,A.S., Pelin,K.,  
Witt,C.C., Bang,M.L., Trombitas,K., Granzier,H., Gregorio,C.C.,



TITLE	Sortmachl,H. and Labelle,S. Identification of muscle specific ring finger proteins as potential regulators of the titin kinase domain
JOURNAL	J. Mol. Biol. 306 (4), 717-726 (2001)
MEDLINE	21140140
PUBMED	11243782
REFERENCE	2 (bases 1 to 2202)
AUTHORS	Centner,T.
TITLE	Direct Submission
JOURNAL	Submitted (12-FEB-2001) Centner T., Structure and Biocomputing, EMBL-Heidelberg, Meyerhofstrasse 1, Heidelberg 69112, Germany
REMARK	Revised by author 22-FEB-2001
FEATURES	Location/Qualifiers
SOURCE	1..2202
misc_feature	/organism="Homo sapiens" /db_xref="taxon:9606" 133..135
gene	/note="putative alternative translation initiation codon" /evidence="not_experimental" 181..1779
CDS	/gene="RNF29" 181..1779 /note="RNF29" /note="alternative" /codon_start=1 /product="ring finger protein 29" /protein_id="C6C32840.1" /db_xref="GI:13160386" /translation="MDNLEKOLICPICLEMEFTKPVVILPCQHNLCRKCSADIFQASNP LYPTRGTTMASGGRFPCSPCRHENVLDHRGVYGLQRMNLVENIIDIYKQESTRPERK SLOPMCEHEHEERINICYLNCCEVPTCSLCYFGAKHDCDOAVPLTHVFORKSLSDGI ALIVGSDNRVQGVISQLEDCTKTEECCRKQKOEICERDVAFLYGLIEERKNEMTOVTT RKQEBLEHVRLIKRSDHLENVSKLVESGIQFMDEPMAYFLQAKTILKKISEAS KAFQMEKIEHGYENMNHFTVNLNREKTIIREIDYREDEDEEBEGEGEGEVG GVAEEVEEVENVQTEPEFEDENPERKASLSQVELQAPALPVSSPEPPALPPADA PYATGVEEVENVQTEPEFEDENPERKASLSQVELQAPALPVSSPEPPALPPADA GQIGPPGSEDSNRYKAAVAAAASRAVSGKETSAPATSQIGFAPPLQGAAPARA SSGASDSEPARHIFESPWLSLNE" join(181..1365,1654..1779) /gene="RNF29" /note="short isoform" /codon_start=1 /product="ring finger protein 29" /protein_id="C6C32839.1" /db_xref="GI:13160385" /translation="MDNLEKOLICPICLEMEFTKPVVILPCQHNLCRKCSADIFQASNP LYPTRGTTMASGGRFPCSPCRHENVLDHRGVYGLQRMNLVENIIDIYKQESTRPERK SLOPMCEHEHEERINICYLNCCEVPTCSLCYFGAKHDCDOAVPLTHVFORKSLSDGI ALIVGSDNRVQGVISQLEDCTKTEECCRKQKOEICERDVAFLYGLIEERKNEMTOVTT RKQEBLEHVRLIKRSDHLENVSKLVESGIQFMDEPMAYFLQAKTILKKISEAS KAFQMEKIEHGYENMNHFTVNLNREKTIIREIDYREDEDEEBEGEGEGEVG GVAEEVEEVENVQTEPEFEDENPERKASLSQVELQAPALPVSSPEPPALPPADA PYATGVEEVEEVENVQTEPEFEDENPERKASLSQVELQAPALPVSSPEPPALPPADA PVTQIGFAPPLQGAAPAAAAGSASDSEPARHIFESPWLSLNE"
BASE COUNT	660 a 489 c 560 g 493 t
ORIGIN	
Query Match	1.9%; Score 27; DB 9; Length 2202;
Best Local Similarity	100.0%; Pied. No. 0.0052;
Matches	27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy	508 ATCATGTGACATCTACAGCAGGAGACTCC 534 
Db	442 ATCATGTGACATCTACAGCAGGAGACTCC 468
RESULT 19	
LOCUS	AK091310 2634 bp mRNA linear PRI 15-JUL-2002
DEFINITION	Homo sapiens cDNA FLJ33991 f1s, clone DPNES2007332, moderately similar to Mus musculus ring-finger protein MURF mRNA.
ACCESSION	AK091310
VERSION	AK091310.1 GI:217499650
KEYWORDS	Oligo capping; f1s (full insert sequence).

SOURCE	Homo sapiens normal dermal fibroblasts (Neonatal Skin) (NHDF2564)
ORGANISM	CDNA to mRNA, clone_1lb:DPNES2 clone:DPNES2007332.
REFERENCE	Homo sapiens
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
TITLE	1 Tashiro,H., Yamazaki,M., Watanabe,K., Kumagai,A., Itakura,S., Fukuzumi,Y., Fujimori,Y., Komiyama,M., Sugiyama,T., Irie,R., Otsuki,T., Sato,H., Wakatsuki,A., Ishii,S., Yamamoto,J., Isono,Y., Kawai-Hio,Y., Saito,K., Nishikawa,T., Kimura,K., Yamashita,H., Matsuo,K., Nakamura,Y., Sekine,M., Kikuchi,H., Kanda,K., Matsumoto,M., Murakawa,K., Kanehori,K., Takahashi-Fujii,A., Oshima,A., Sugiyama,A., Kawakami,B., Suzuki,Y., Sugano,S., Nagahari,K., Masuhara,Y., Nagai,K. and Isogai,T.
JOURNAL	NEDO human cDNA sequencing project
REFERENCE	Unpublished
AUTHORS	2 (bases 1 to 2634)
TITLE	Isogai,T. and Yamamoto,J.
JOURNAL	Direct Submission
COMMENT	Submitted (04-JUL-2002) Takao Isogai, FLJ Project(HRI Team); 2-6-7 Kazuo-Kamachi, Kisarazu, Chiba 292-0812, Japan (E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986) NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; CDNA full insert sequencing: Research Association for Biotechnology (RAB); cDNA library construction: Helix Research Institute (HRI) (supported by Japan Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB, HRI, and Biotechnology Center, National Institute of Technology and Evaluation; clone selection for full insert sequencing: HRI and RAB; annotation: HRI and RAB.
FEATURES	Location/Qualifiers
SOURCE	1..2634
	/organism="Homo sapiens"
	/db_xref="taxon:9606"
	/clone="DPNES2007332"
	/cell_type="normal dermal fibroblasts (Neonatal Skin)"
	(NHDF2564)"
	/clone_1lb="DPNES2"
	/note="cloning vector: PME18SFL3-primary culture, normal dermal fibroblasts"
BASE COUNT	793 a 580 c 586 g 675 t
ORIGIN	
Query Match	1.9%; Score 27; DB 9; Length 2634;
Best Local Similarity	100.0%; Prid. No. 0.0052;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
OY	508 ATCATTTGACATCTTACAAGCAGAGATCC 534       Db 686 ATCATTGACATCTTACAAGCAGAGATCC 712
RESULT 20	
AC090195	121721 bp DNA linear PRI 29-JUN-2002
LOCUS	
DEFINITION	Homo sapiens chromosome 8, clone RP11-366K18, complete sequence.
ACCESSION	AC090195
VERSION	AC090195.5 GI:21629404
KEYWORDS	HTG.
SOURCE	human.
ORGANISM	Homo sapiens
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE	1 (bases 1 to 121721)
AUTHORS	Birtten,B., Nusbaum,C. and Lander,E.
TITLE	Unpublished
JOURNAL	2 (bases 1 to 121721)
REFERENCE	Birtten,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,S., Barta,N., Bastien,V., Boguslavsky,L., Boukhalter,B., Brown,A., Camarero,J., Campopiano,A., Choepel,Y., Colangelo,M., Collins,S., Collamore,A., Cooke,P., D'Arcangelo,K., Dewar,K., Diaz,J.S., Dodge,S., Fato,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J.,



Gardyna, S., Ginde, S., Goyette, M., Graham, L., Grand-Pierre, N., Hagos, B., Heatford, A., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Karatas, A., Laroque, K., Lamazares, R., Landers, T., Lenoczky, D., Levine, R., Liu, G., Maclean, C., Macdonald, P., Marcquis, N., Matthews, C., McCarthy, M., McEwan, P., McKernan, K., Mcneeters, R., Meldrim, J., Menus, L., Mihova, T., Menga, V., Murphy, T., Naylor, J., Nguyen, C., Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R., Riback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Roselli, M., Roy, A., Santos, R., Schauer, S., Schupback, R., Seaman, S., Severy, P., Sougne, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Strauss, N., Subramanian, A., Talamas, J., Testaye, S., Theodore, J., Travers, M., Travis, N., Triggillo, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
Submitted (17-FEB-2001) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA  
3 (bases 1 to 121721)

Anderson, S., Baran, N., Bastien, V., Bloom, T., Boguslavsky, L., Boukhaltier, B., Brown, A., Camarata, J., Campopiano, A., Chang, J., Chazaro, B., Choepel, Y., Colangelo, M., Collins, S., Collymore, A., Cook, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J.S., Dodge, S., Fato, S., Ferreira, P., Fitzgerald, M., FitzHugh, W., Gage, D., Galagan, J., Gardyna, S., Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N., Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Laroque, K., Lamazares, R., Landers, T., Lenoczky, J., Levine, R., Lindblad-Toh, K., Liu, G., Maclean, C., Macdonald, P., McEwan, P., McKernan, K., Meldrim, J., Menus, L., Mihova, T., Mienna, C.H., Murphy, T., Naylor, J., Nguyen, C., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R., Riback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Roselli, M., Roy, A., Santos, R., Schauer, S., Schupback, R., Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Strauss, N., Subramanian, A., Talamas, J., Testaye, S., Theodore, J., Topham, K., Travers, M., Travis, N., Triggillo, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
Direct Submission  
Submitted (26-JUN-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA  
4 (bases 1 to 121721)

Baran, N., Bastien, V., Bloom, T., Boguslavsky, L., Boukhaltier, B., Camarata, J., Chang, J., Chazaro, B., Choepel, Y., Collymore, A., Cook, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J.S., Dodge, S., Gardyna, S., Gord, S., Graham, L., Grand-Pierre, N., Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, G., Maclean, C., Macdonald, P., Major, J., Matthews, C., McCarthy, M., Meldrim, J., Menus, L., Mihova, T., Mienna, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Roy, A., Schauer, S., Schupback, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Talamas, J., Testaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

TITLE  
JOURNAL  
COMMENT  
Submitted (29-JUN-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA  
On Jun 29, 2002 this sequence version replaced g1.21592056.  
All repeats were identified using RepeatMasker:  
Smit, A.F.A. & Green, P. (1996-1997)  
http://ftp.genome.washington.edu/RM/RepeatMasker.html

Center: Whitehead Institute/ MIT Center for Genome Research

# FEATURES

## source

Center code: WIGR  
Web site: <http://www.seq.wi.mit.edu>  
Contact: [sequence\\_submissions@genome.wi.mit.edu](mailto:sequence_submissions@genome.wi.mit.edu)  
Project Information  
Center project name: LI2201  
Center clone name: 366\_K\_18  
-----  
The first 121.7 kb of this clone are being submitted.  
The remainder overlaps accession number AC090196 [WIGR project LI2201].

## Location/Qualifiers

1..121721	Location/Qualifiers
/organism="Homo sapiens"	
/db_xref="taxon:9606"	
/chromosome="8"	
/map="8"	
/clone="RP11-366K18"	
/clone_lib="RPC1-11 Human Male BAC"	
470..519	
/rpt_family="(TG)n"	repeat_region
624..842	
/rpt_family="L2"	repeat_region
complement(1206..1457)	
/rpt_family="MER34"	repeat_region
complement(1470..2979)	
/rpt_family="T199GT2"	repeat_region
complement(2980..3164)	
/rpt_family="MER02"	repeat_region
3264..3380	
/rpt_family="MER02"	repeat_region
3707..3842	
/rpt_family="MIR"	repeat_region
complement(3937..4400)	
/rpt_family="WLT1B"	repeat_region
4253..4258	
/note="<30 qual SINGL region"	repeat_region
4310..4314	
/note="<30 qual SINGL region"	repeat_region
4338..4384	
/note="<30 qual SINGL region"	repeat_region
4950..5208	
/note="<30 qual SINGL region"	repeat_region
/rpt_family="LTR16A"	repeat_region
complement(5263..5702)	
/rpt_family="HAL1"	repeat_region
complement(5703..5736)	
/rpt_family="Alu0"	repeat_region
5737..5776	
/rpt_family="(TG)n"	repeat_region
complement(5777..6061)	
/rpt_family="Alu0"	repeat_region
complement(6062..6428)	
/rpt_family="HAL1"	repeat_region
7448..7707	
/rpt_family="Alu0"	repeat_region
complement(8751..8825)	
/rpt_family="trna-Arg-CCC"	repeat_region
9180..9272	
/rpt_family="trna-Tyr-TAC"	repeat_region
9801..9889	
/rpt_family="trna-Tyr-TAC"	repeat_region
10004..10076	
/rpt_family="trna-Ala-GCA"	repeat_region
11489..11722	
/rpt_family="MIR"	repeat_region
11944..12107	
/rpt_family="L1ME4A"	repeat_region
12145..12450	
/rpt_family="Alu0"	repeat_region
12492..13165	
/rpt_family="L1M4"	repeat_region
13167..13314	
/rpt_family="Charlie5"	repeat_region

```
unsure      complement(13332..13337)
             /note="<30 qual SNGL region"
repeat_region complement(13347..13464)
             /rpt_family="I2"
unsure      complement(13348..13352)
             /note="<30 qual SNGL region"
unsure      complement(13382..13387)
             /note="<30 qual SNGL region"
unsure      complement(13407..13421)
             /note="<30 qual SNGL region"
unsure      complement(13459..13473)
             /note="<30 qual SNGL region"
unsure      complement(13459..13473)
             /note="<30 qual SNGL region"
repeat_region repeat_region
             /rpt_family="AluY"
             14684..14744
             /rpt_family="(TGA)n"
repeat_region repeat_region
             complement(15595..15784)
             /rpt_family="MER3"
             15805..15943
repeat_region repeat_region
             /rpt_family="MIR"
             16103..16825
             /rpt_family="MER21B"
repeat_region repeat_region
             17275..17674
             /rpt_family="MER57A"
repeat_region repeat_region
             17956..18054
             /rpt_family="MER4A"
repeat_region repeat_region
             18055..18414
             /rpt_family="MER4A"
repeat_region repeat_region
             18417..19264
             /rpt_family="MER4D"
repeat_region repeat_region
             19358..19402
             /rpt_family="GA-rich"
repeat_region repeat_region
             19411..19683
             /rpt_family="AluJo"
repeat_region repeat_region
             complement(19738..19867)
```

Query Match 1.9% Score 27; DB 9; Length 121721;

Best Local Similarity 100.0%; Pred. No. 0.0058; Mismatches 0; Indels 0; Gaps 0;

Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 508 ATCATGTGACATCTACAGCAGAGTCC 534

|||||

Db 24260 ATCATGTGACATCTACAGCAGAGTCC 24286

Search completed: December 3, 2002, 12:09:29  
Job time : 5217 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: December 3, 2002, 13:10:09 ; Search time 60 seconds  
(without alignments)  
2349.204 Million cell updates/sec

Title: US-09-908-988B-2  
Perfect score: 1912  
Sequence: 1 MNFTVGRKPLIGAHNMNDL.....EGNAGLEERLDVPEGSGIH 366

Scoring table:  
BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 6.0, Ygapext 7.0  
Delop 6.0, Delext 7.0

Searched: 341543 segs, 192557720 residues  
Total number of hits satisfying chosen parameters: 683086

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame\_plus\_p2n\_model -DEV=xlh  
Q=/cgn2\_1/USPFO.spool/US09908988/runat\_26112002.112156.11295/app.query.fasta.1.519  
-DB=Published.Applications\_NA -OFMT=fastap -SUFFIX=rmpb -MINMATCH=0.1  
-LOOPCL=0 -LOOEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blomsu62  
-TRANS=Numan40.cdi -LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100  
-THR MIN=0 -ALIGN=20 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0  
-MAXLEN=2000000000 -USER=US0908988.ecgn.1.1.21 -runat\_26112002.112156.11295  
-NCPU=6 -ICPU=3 -NO\_XIPYX -NO\_MAP -LARGEDUTRY -NEG\_SCORES=0 -WAIT -LONGLOG  
-DEV.TIMEOUT=120 -WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Published Applications\_NA: \*  
1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq: \*  
2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq: \*  
3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq: \*  
4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq: \*  
5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq: \*  
6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq: \*  
7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq: \*  
8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq: \*  
9: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq: \*  
10: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq: \*  
11: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq: \*  
12: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq: \*  
13: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq: \*  
14: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq: \*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1912	100.0	1431	10	US-09-908-988B-1
2	1162	60.8	2590	10	US-09-908-988B-3
3	1147.5	60.0	1762	10	US-09-764-864-34
4	1110	58.1	1597	10	US-09-908-988B-5

*His Appl No 162 not published*  
*At No 913*  
*is listed*

5	1091.5	57.1	1781	10	US-09-764-864-21	Sequence 21, Appl
6	614	32.1	587	10	US-09-764-864-493	Sequence 493, App
7	416	21.8	446	10	US-09-764-864-482	Sequence 482, App
8	280	14.6	1039	10	US-09-880-192-25	Sequence 25, Appl
9	274	14.3	1755	10	US-09-925-300-729	Sequence 729, App
10	234	12.2	2981	10	US-09-925-301-101	Sequence 101, App
11	225	11.8	599	10	US-09-764-864-1594	Sequence 1594, Ap
12	225	11.8	599	10	US-09-764-864-1595	Sequence 1595, Ap
13	225	11.8	599	10	US-09-764-877-2549	Sequence 2549, Ap
14	225	11.8	599	10	US-09-764-877-2550	Sequence 2550, Ap
15	225	11.8	625	10	US-09-764-864-46	Sequence 46, Appl
16	225	11.7	625	10	US-09-764-877-269	Sequence 269, App
17	223.5	11.7	7596	10	US-09-728-952-1	Sequence 1, Appl1
18	200.5	10.5	1332	10	US-09-250-883-14	Sequence 14, Appl
19	199	10.4	3441	10	US-09-880-192-41	Sequence 41, Appl
20	196	10.3	3826	10	US-09-927-091-3	Sequence 3, Appl1
21	191.5	10.0	479	10	US-09-920-300A-1569	Sequence 1569, Ap
22	191.5	10.0	479	12	US-10-033-528-1569	Sequence 1569, Ap
23	186	9.7	3033	12	US-10-044-090-508	Sequence 508, App
24	174.5	9.1	1400	10	US-09-764-864-56	Sequence 56, Appl
25	161	8.4	774	10	US-09-764-864-40	Sequence 40, Appl
26	156	8.2	459	10	US-09-864-761-31683	Sequence 31683, A
27	152	7.9	737	10	US-09-910-943-471	Sequence 471, App
28	151.5	7.9	3035	10	US-09-864-864-311	Sequence 311, Appl
29	143.5	7.5	560	10	US-09-864-761-1296	Sequence 1296, Ap
30	143.5	7.5	560	10	US-09-864-761-26672	Sequence 26672, A
31	139	7.3	2089	10	US-09-764-864-1598	Sequence 1598, Ap
32	136	7.1	32190	10	US-09-764-878-201	Sequence 201, App
33	135	7.1	45845	10	US-09-927-091-6	Sequence 6, Appl1
34	134.5	7.0	456	10	US-09-864-761-1296	Sequence 1296, Ap
35	134.5	7.0	456	10	US-09-864-761-10030	Sequence 10030, A
36	134	7.0	488	10	US-09-864-761-15156	Sequence 15156, A
37	133	7.0	32249	10	US-09-764-878-202	Sequence 202, App
38	132	6.9	3587	10	US-09-862-436-294	Sequence 294, App
39	132	6.9	3587	10	US-09-880-107-2401	Sequence 2401, App
40	131	6.9	49744	10	US-09-927-091-4	Sequence 4, Appl1
41	130.5	6.8	262	10	US-09-250-883-1	Sequence 1, Appl1
42	130.5	6.8	1310	10	US-09-764-864-1651	Sequence 1651, Ap
43	130	6.8	6421	10	US-09-764-864-1650	Sequence 1650, Ap
44	127	6.6	429	10	US-09-864-761-18056	Sequence 18056, A
45	127	6.6	429	10	US-09-864-761-18059	Sequence 18059, A

#### ALIGNMENTS

RESULT 1  
US-09-908-988B-1  
: Sequence 1, Application US/09908988B  
: Patent No. US20020127690A1  
: GENERAL INFORMATION:  
: APPLICANT: OLSON, ERIC  
: APPLICANT: SPENCER, JEFFREY A.  
: TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR STABILIZING MICROTUBULES  
: FILE REFERENCE: MYOG:0280US  
: CURRENT APPLICATION NUMBER: US/09/908, 988B  
: PRIOR FILING DATE: 2000-07-18  
: PRIOR APPLICATION NUMBER: 60/219, 020  
: NUMBER OF SEQ ID NOS: 6  
: SOFTWARE: PatentIn Ver. 2.1  
: SEQ ID NO 1  
: LENGTH: 1431  
: TYPE: DNA  
: ORGANISM: Mus musculus  
: FEATURE:  
: NAME/KEY: CDS  
: LOCATION: (199)..(1296)  
US-09-908-988B-1  
Alignment Scores: 1.51e-210 Length: 1431  
Pred. No.: 1912.00 Matches: 366

Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 10 Gaps: 0

US-09-908-988B-2 (1-366) x US-09-908-988B-1 (1-1431)

QY 1 MetaspheThValGlyPheLysProLeuLeuGlyAspAlaHisAsnMetAspAsnLeu 20  
DB 199 ATGAACTTCACGGTGGGTTTCAAGCCGCTCTAGGGGATGGCCAAACATGACAACTTG 258  
QY 21 GlnLysGlnLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 40  
DB 259 GAGAGCAGCTCATTTGGCCCCATCTCGCTGAGATGTTCTCAAGCCCGGTGGATCTTG 318  
QY 41 ProCysGlnHisAsnLeuLysArgLysCysAlaAsnAspValPheGlnAlaSerAsnPro 60  
DB 319 CCGTCCACACACACCTGTGGCCGAGGTGGCCACAGCATCTTCCAGGCTCTTAATCT 378  
QY 61 LeuTrpGlnSerArgGlySerThrThrValSerSerGlyGlyArgPheArgCysProSer 80  
DB 379 CTGTGGCAATCCCGGGGCTCCACACGGGTCTTCAAGAGAGCTTCCGATGCCATCT 438  
QY 81 CysArgHisGlyValValLeuAspArgHisGlyValTyrGlyLeuGlnArgAsnLeuLeu 100  
DB 439 TGTAGGCACAGAGTGTCTCGACAGCATGTGTCTATGAGCTCGACGGACCTGCTA 498  
QY 101 ValGlnAsnIleIleAspIleTyrLysGlnGlnLysSerArgProLeuHisAlaLysAla 120  
DB 499 GTGGAGAACATATTGTACATTCACAAAGAGAGAGTCTCCGGCCACATGCACGCAAGGCT 558  
QY 121 GlnGlnHisLeuMetCysGlnGlnHisGlyAspGlnLysIleAsnIleTyrCysLeuSer 140  
DB 559 GACACGACACCTCATGTGTGACAGCAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 618  
QY 141 CysGlnValProThrCysSerLeuCysLysValPheGlyAlaHisLysAspCysGlnVal 160  
DB 619 TGGCGAGTGGCCACCTGCTCTCTGCAAGGTTTGGCGCCCCACAGAGACTGTGAGGTG 678  
QY 161 AlaProLeuProThrIleTyrLysArgGlnLysSerGlnLeuSerArgGlyIleAlaMet 180  
DB 679 GCCCTCTGGCCCATTTACAAACGCCAGAAAGAGTGAAGTGAAGTGAAGTGAAGTGAAG 738  
QY 181 LeuValAlaGlyAsnAspArgValGlnAlaValIleThrGlnMetGlnGlnValCysGln 200  
DB 739 CTGGTGGGGGCAATGACCGTGTGCGAGCATATCACCCAGATGGAGAGGTGTGCCAG 798  
QY 201 ThrIleGlnAspAsnSerArgArgGlnLysGlnLeuLeuAsnGlnArgPheGlnThrLeu 220  
DB 799 ACCATTTGAGGACAAACAGCCGACAGAGAGCAACTGTTAAACAGAGGTTGAGACCTCG 858  
QY 221 CysAlaValLeuGlnGlnArgLysGlyGlnLeuLeuGlnAlaLeuAlaArgGlnGln 240  
DB 859 TGGCGGATTTTGGAGAGGCGCAAGGCGAACTGTTCAACACACTGGCCCGGAGACAGAG 918  
QY 241 GlnLysLeuGlnArgValArgGlyLeuIleArgGlnTyrGlnLysPheHisLeuGlnLys 260  
DB 919 GAGAAAGTTGACGCGGTGGGGGCTCATCCGCAAGAGAGAGAGAGAGAGAGAGAGAGAG 978  
QY 261 SerLysLeuValGlnSerAlaIleGlnSerMetGlnGlnProGlnMetAlaLeuTyrLeu 280  
DB 979 TCAAAAGCTGGTAGTCCGCCCATCCAGTCCATGGAGAGAGCGGAGAGGCTCTTACCTC 1038  
QY 281 GlnGlnAlaLysGlnLeuIleAsnLysValGlyAlaMetSerLysValGlnLeuAlaGly 300  
DB 1039 CAGCGGCAAAAGAGCTGTATCAACAGGTCGGGGCAATGTCGAAGGTGGAGCTGGCAGGA 1098  
QY 301 ArgProGlnProGlnTyrGlnSerMetGlnInPheSerValSerValGlnHisValAla 320  
DB 1099 CGGCCGGAGCCAGGCTATGAGAGCATGGAATTTCTGTGTGAGCGGTGGAGCAGCTGGCC 1158  
QY 321 GlnMetLeuArgThrIleAspPheGlnProGlyAlaAlaGlnLysArgGlnLysAspAsp 340

DB 1159 GAAATGTTGGCAACATGACTTCAGCCGGCGCGCTGGGGATGAAGAGATGACAGC 1218  
QY 341 MetAlaLeuAspArgGlnGlnGlyAlaAlaGlyLeuGlnGlnArgLeuAspValPro 360  
DB 1219 ATGGCTTTGGATGGGAGAGGCGCAATGCGGGCTGGAGAGAGAGCGGCTGGACGTGCA 1278  
QY 361 GlnGlySerGlyLeuHis 366  
DB 1279 GAGGCTCAGCGCTGCAC 1296

RESULT 2  
US-09-908-988B-3  
; Sequence 3, Application US/0908988B  
; Patent No. US20020127690A1  
; GENERAL INFORMATION:  
; APPLICANT: OLSON, ERIC  
; APPLICANT: SPENCER, JEFFREY A.  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR STABILIZING MICROTUBULES  
; FILE REFERENCE: MYOG:028US  
; CURRENT APPLICATION NUMBER: US/09/908, 988B  
; PRIOR FILING DATE: 2000-07-18  
; NUMBER OF SEQ ID NOS: 6  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 3  
; LENGTH: 2590  
; TYPE: DNA  
; ORGANISM: Mus musculus  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (80)..(1714)  
US-09-908-988B-3

Alignment Scores:  
Pred. No.: 7,37e-124 Length: 2590  
Score: 1162.00 Matches: 226  
Percent Similarity: 76.99% Conservative: 55  
Best Local Similarity: 61.92% Mismatches: 76  
Query Match: 60.77% Indels: 8  
Gaps: 4

US-09-908-988B-2 (1-366) x US-09-908-988B-3 (1-2590)

QY 1 MetaspheThValGlyPheLysProLeuLeuGlyAspAlaHisAsnMetAspAsnLeu 20  
DB 80 ATGAGCACTTCTGTGAATTCAAACTTTTCTCAAAAGAGCAGACCATGGATTAATTG 139  
QY 21 GlnLysGlnLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 40  
DB 140 GAAAGCAACTGTATCTCCATCTGCTAGAGAGTTCAGAGAACTGTGGTCAATTC 199  
QY 41 ProCysGlnHisAsnLeuLysArgLysCysAlaAsnAspValPheGlnAlaSerAsnPro 60  
DB 200 CTTTGGCAGACAAACCTGTGACAGAAATGTGCCAGTGCATCTTCCAGGCTCTTAACCG 259  
QY 61 LeuTrpGlnSerArgGlySerThrThrValSerSerGlyGlyArgPheArgCysProSer 80  
DB 260 TACTTACCCACAG 319  
QY 81 CysArgHisGlyValValLeuAspArgHisGlyValTyrGlyLeuGlnArgAsnLeuLeu 100  
DB 320 TGCAGACATGAGGTGGTGTAGACAGACATGGGGTCTATGGACTGCAAGAGAACTGCTC 379  
QY 101 ValGlnAsnIleIleAspIleTyrLysGlnGlnLysSerArgProLeuHisAlaLysAla 120  
DB 380 GTGGAAACATTTTGTATCTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 436  
QY 121 GlnGlnHisLeuMetCysGlnGlnHisGlyAspGlnLysIleAsnIleTyrCysLeuSer 140  
DB 437 GACACGAGCC--ATGTGTGAAGAGACATGAAAGAGAAAGAGAGAGAGAGAGAGAGAGAG 493



Db 1153 T-----GAAGAAGAGAGAGCGCGAGAGAGAGAA----- 1186  
QY 360 oGluGlyserGly 364  
Db 1187 -GAAGAGAGAGCA 1198  
RESULT 4  
US-09-908-988b-5  
; Sequence 5, Application US/09908988B  
; Patent No. US20020127690A1  
; GENERAL INFORMATION:  
; APPLICANT: OLSON, ERIC  
; APPLICANT: SPENCER, JEFFREY A.  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR STABILIZING MICROTUBULES  
; FILE REFERENCE: MYOG:02805  
; CURRENT APPLICATION NUMBER: US/09/908,988B  
; PRIOR FILING DATE: 2000-07-18  
; PRIOR APPLICATION NUMBER: 60/219,020  
; NUMBER OF SEQ ID NOS: 6  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 5  
; LENGTH: 1597  
; TYPE: DNA  
; ORGANISM: Mus musculus  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (299)..(1327)  
US-09-908-988b-5  
Alignment Scores:  
Pred. No.: 3,41e-118 Length: 1597  
Score: 1110.00 Matches: 212  
Percent Similarity: 78.53% Conservative: 55  
Best Local Similarity: 62.35% Mismatches: 63  
Query Match: 58.05% Indels: 10  
DB: 10 Gaps: 3  
US-09-908-988b-2 (1-366) x US-09-908-988b-5 (1-1597)  
QY 4 ThrValGlyPheLys-----ProLeuGluGlyAspAlaHisasn 16  
Db 278 ACAGTCCGATTTCAAGCAATATGATTAATATAGCTGATCTCTGATGAAAGCGCT 337  
QY 17 MetAspAsnLeuGluGlyGlnLeuIleCysProIleCysLeuGluMetPheSerLysPro 36  
Db 338 ATGGAGAACCTGGAGAGAGAGCTGATCTGCCCATCTGCTGGAGATGTTTACCAAGCGCT 397  
QY 37 ValValIleLeuProCysGlnHisAsnLeuCysArgLysCysAlaAspValPheGln 56  
Db 398 GTGGTCAATCTGCGCTGCCACACACACCTCTGCCGAGAGTGTGCCAAGACATCTTCAG 457  
QY 57 AlaSerAsnProLeuTrpGlnSerArgLysSerThrValSerSerGlyArgPhe 76  
Db 458 GCTGGAGATTCCTACTGACCAACCGCGGTGCTCACTGTCATGCTGGAGTTCGCTTC 517  
QY 77 ArgCysProSerCysArgHisGluValValLeuAspArgHisGlyValIleGlyLeuGln 96  
Db 518 CGTGGCCCTCTGCGCGCATGAGATGATGACCGGCGGAGTGTACGGCCCTGAG 577  
QY 97 ArgAsnLeuLeuValGluAsnIleIleAspIleTyrLysGlnGlu---SerSerArgPro 115  
Db 578 AGGAACCTGCTGCTGAGAAACATCATTTACACAGAGAGAGTCTCCAGTCGGCCC 637  
QY 116 LeuHisAlaLysAlaGluGlnHisLeuMetCysGluGluHisGluAspGluLysIleasn 135  
Db 638 CTG-----CAAGAGGAGAGCCAGCGGATGTGCAAGAGAACAGAGAGATCAAC 691  
QY 136 IleTyrCysLeuSerCysGluValProThrCysSerLeuCysLysValPheGlyAlaHis 155  
Db 692 ATCTACTGTCTCAGCTGAGTGTGCTACTGTCTCTTGTGCAAGTGTGTTGGGGCTCAC 751

QY 156 LysAspCysGluValAlaProLeuProThrIleTyrLysArgGlnLysSerGluLeuSer 175  
Db 752 CAGGCTGTGAGGTGCCCTTGGCAAGACATCTTCCAGAGACAGATGAGCTGAGT 811  
QY 176 AspGlyIleAlaMetLeuValAlaGlyAsnAspArgValGlnAlaValIleThrGlnMet 195  
Db 812 AACGTGATCTCCATGCTGTGGTGGGGAGACCGAGCGAGTGCAGAGATCATCTCAGCTG 871  
QY 196 GluGluValCysGlnThrIleGluAspAsnSerArgArgGlnLysGlnLeuLeuAsnGln 215  
Db 872 GAGACTGCTGAGAGTGCACCAAGAGAAATAGCCACAGGAGAGAGAGAGAGCTGAGT 931  
QY 216 ArgPheGluThrLeuCysAlaValIleGluGluGluArgLysGlyGluLeuGlnAlaLeu 235  
Db 932 AAGTTTGACACCTCTTACGCGCATCTGTGATGAGAGAGAGAGAGAGAGAGAGAGAG 991  
QY 236 AlaArgGluGlnGluGluLysLeuGlnArgValArgGlyLeuIleArgGlnIleGlyLys 255  
Db 992 AGCAG 1051  
QY 256 HisLeuGluGlySerSerLysLeuValGluSerAlaIleGlnSerMetGluLysProGln 275  
Db 1052 CAGCTGGAAAGTCCACCAAGCTTGTGAGAGACCCCATCTCCAGTGTGAGAGAGAG 1111  
QY 276 MetAlaLeuTyrLeuGlnAlaLysGlnLeuIleAsnLysValGlyAlaMetSerLys 295  
Db 1112 GGGGCTACCTCTCTTCAAGTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1171  
QY 296 ValGluLeuAlaGlyArgProGluProGlyTyrGlnSerMetGluGlnPheSerValSer 315  
Db 1172 GCGTGCAGCTGGGAG 1231  
QY 316 ValGluHisValAlaGluMetLeuArgThrIleAspPheGlnProGlyAlaAlaLysP 335  
Db 1232 TTAGAACATATACAGAGAGGCTTGTGAGGCGCATGACTTGGGAGAGAGATGTAT 1291  
RESULT 5  
US-09-764-864-21  
; Sequence 21, Application US/09764864  
; Patent No. US20020132753A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: PT223  
; CURRENT APPLICATION NUMBER: US/09/764,864  
; PRIOR FILING DATE: 2001-01-17  
; PRIOR APPLICATION data removed - consult PALM or file wrapper  
; NUMBER OF SEQ ID NOS: 1792  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 21  
; LENGTH: 1781  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-764-864-21  
Alignment Scores:  
Pred. No.: 5,41e-116 Length: 1781  
Score: 1091.50 Matches: 209  
Percent Similarity: 77.71% Conservative: 63  
Best Local Similarity: 59.71% Mismatches: 67  
Query Match: 57.09% Indels: 11  
DB: 10 Gaps: 4  
US-09-908-988b-2 (1-366) x US-09-764-864-21 (1-1781)  
QY 10 LeuLeuGlyLysAlaHisAsnMetAspAsnLeuGluLysGlnLeuIleCysProIleCys 29  
Db 137 CTGATCCAGATGGAGATCCCATGAGAACTTGGAGAGAGAGAGAGAGAGAGAGAGAG 196  
QY 30 LeuGluMetPheSerLysProValValIleLeuProCysGlnHisAsnLeuCysArgLys 49  
Db 197 CTGAGATGTTTACCAAGCAGAGTGTATCTTGTCCGTCAGACAGACACCTGTGCGGAG 256



OS-09-908-988B-2 (1-300) X OS-09-880-192-23 (1-1035)

5  
 6  
 7  
 8  
 9  
 10  
 11  
 12  
 13  
 14  
 15  
 16  
 17  
 18  
 19  
 20  
 21  
 22  
 23  
 24  
 25  
 26  
 27  
 28  
 29  
 30  
 31  
 32  
 33  
 34  
 35  
 36  
 37  
 38  
 39  
 40  
 41  
 42  
 43  
 44  
 45  
 46  
 47  
 48  
 49  
 50  
 51  
 52  
 53  
 54  
 55  
 56  
 57  
 58  
 59  
 60  
 61  
 62  
 63  
 64  
 65  
 66  
 67  
 68  
 69  
 70  
 71  
 72  
 73  
 74  
 75  
 76  
 77  
 78  
 79  
 80  
 81  
 82  
 83  
 84  
 85  
 86  
 87  
 88  
 89  
 90  
 91  
 92  
 93  
 94  
 95  
 96  
 97  
 98  
 99  
 100  
 101  
 102  
 103  
 104  
 105  
 106  
 107  
 108  
 109  
 110  
 111  
 112  
 113  
 114  
 115  
 116  
 117  
 118  
 119  
 120  
 121  
 122  
 123  
 124  
 125  
 126  
 127  
 128  
 129  
 130  
 131  
 132  
 133  
 134  
 135  
 136  
 137  
 138  
 139  
 140  
 141  
 142  
 143  
 144  
 145  
 146  
 147  
 148  
 149  
 150  
 151  
 152  
 153  
 154  
 155  
 156  
 157  
 158  
 159  
 160  
 161  
 162  
 163  
 164  
 165  
 166  
 167  
 168  
 169  
 170  
 171  
 172  
 173  
 174  
 175  
 176  
 177  
 178  
 179  
 180  
 181  
 182  
 183  
 184  
 185  
 186  
 187  
 188  
 189  
 190  
 191  
 192  
 193  
 194  
 195  
 196  
 197  
 198  
 199  
 200  
 201  
 202  
 203  
 204  
 205  
 206  
 207  
 208  
 209  
 210  
 211  
 212  
 213  
 214  
 215  
 216  
 217  
 218  
 219  
 220  
 221  
 222  
 223  
 224  
 225  
 226  
 227  
 228  
 229  
 230  
 231  
 232  
 233  
 234  
 235  
 236  
 237  
 238  
 239  
 240  
 241  
 242  
 243  
 244  
 245  
 246  
 247  
 248  
 249  
 250  
 251  
 252  
 253  
 254  
 255  
 256  
 257  
 258  
 259  
 260  
 261  
 262  
 263  
 264  
 265  
 266  
 267  
 268  
 269  
 270  
 271  
 272  
 273  
 274  
 275  
 276  
 277  
 278  
 279  
 280  
 281  
 282  
 283  
 284  
 285  
 286  
 287  
 288  
 289  
 290  
 291  
 292  
 293  
 294  
 295  
 296  
 297  
 298  
 299  
 300  
 301  
 302  
 303  
 304  
 305  
 306  
 307  
 308  
 309  
 310  
 311  
 312  
 313  
 314  
 315  
 316  
 317  
 318  
 319  
 320  
 321  
 322  
 323  
 324  
 325  
 326  
 327  
 328  
 329  
 330  
 331  
 332  
 333  
 334  
 335  
 336  
 337  
 338  
 339  
 340  
 341  
 342  
 343  
 344  
 345  
 346  
 347  
 348  
 349  
 350  
 351  
 352  
 353  
 354  
 355  
 356  
 357  
 358  
 359  
 360  
 361  
 362  
 363  
 364  
 365  
 366  
 367  
 368  
 369  
 370  
 371  
 372  
 373  
 374  
 375  
 376  
 377  
 378  
 379  
 380  
 381  
 382  
 383  
 384  
 385  
 386  
 387  
 388  
 389  
 390  
 391  
 392  
 393  
 394  
 395  
 396  
 397  
 398  
 399  
 400  
 401  
 402  
 403  
 404  
 405  
 406  
 407  
 408  
 409  
 410  
 411  
 412  
 413  
 414  
 415  
 416  
 417  
 418  
 419  
 420  
 421  
 422  
 423  
 424  
 425  
 426  
 427  
 428  
 429  
 430  
 431  
 432  
 433  
 434  
 435  
 436  
 437  
 438  
 439  
 440  
 441  
 442  
 443  
 444  
 445  
 446  
 447  
 448  
 449  
 450  
 451  
 452  
 453  
 454  
 455  
 456  
 457  
 458  
 459  
 460  
 461  
 462  
 463  
 464  
 465  
 466  
 467  
 468  
 469  
 470  
 471  
 472  
 473  
 474  
 475  
 476  
 477  
 478  
 479  
 480  
 481  
 482  
 483  
 484  
 485  
 486  
 487  
 488  
 489  
 490  
 491  
 492  
 493  
 494  
 495  
 496  
 497  
 498  
 499  
 500  
 501  
 502  
 503  
 504  
 505  
 506  
 507  
 508  
 509  
 510  
 511  
 512  
 513  
 514  
 515  
 516  
 517  
 518  
 519  
 520  
 521  
 522  
 523  
 524  
 525  
 526  
 527  
 528





```

Db 900 CAGCTGATCACTCTTAAAGAGCATGATATCGCTCTCGCCCTTGAGAGCTA 959
Oy 239 GlnGluGluValLeuGlnArgValArgGlyLeuIleArgGlnIleGlyAspHisLeuGlu 258
Db 960 GACCTGGCCATTCACAAATGCAATGAGTGGCCATCACCAGCTTCTTGCAACATCTCC 1019
Oy 259 GlySerSerIleuValGluSerAlaIleGlnSerMetGluGluProGlnMetAlaLeu 278
Db 1020 CACCTCAGCAGCGCTG-----ATCGCTCAGCTAAGAGAGACGACAG----- 1061
Oy 279 TyrLeuGlnGlnAlaLeuGlnLeuIleAsnLysValGly---AlaMetSerLysValGlu 297
Db 1062 -----CAGCCACACAGAGAGCTCTCGCAGAGCATTTGGGACACATTCAGAGCGCTGAA 1115
Oy 298 LeuAlaGlyArgProGluProGlyTyr----- 306
Db 1116 AGAATCAGGATTCCTGAACTTGAGTCACTCCAGATTTCGACAGAAATCCACATT 1175
Oy 307 -----GluSerMetGluGlnPheSerValSerValGlu 317
Db 1176 TTGGCCAAAATGCTATCTTCTGAGGAGAGCTTAAGCAGATTTCACAGAAATATCGAG 1235
Oy 318 HisValAlaGluMetLeuArgThrIle 326
Db 1236 TCAGATATGAGAGAAATCCAGAAATTA 1262

RESULT 11
US-09-764-864-1594/c
; Sequence 1594, Application US/09764864
; Patent No. US20020132753A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT223
; CURRENT APPLICATION NUMBER: US/09/764,864
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1792
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1594
; LENGTH: 599
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-864-1594

Alignment Scores:
Pred. No.: 5 23e-17 Length: 599
Score: 225.00 Matches: 59
Percent Similarity: 43.52% Conservative: 35
Best Local Similarity: 27.31% Mismatches: 98
Query Match: 11.77% Indels: 24
DB: 10 Gaps: 6

US-09-908-988b-2 (1-366) x US-09-764-864-1594 (1-599)
Oy 17 MetAspAsnLeuGluGlnLeuIleCysProIleCysLeuGluMetPheSerLysPro 36
Db 595 ATGCACAAATTTTGAGAGAGCTTAAGTATGTTATGTTTGAAGATTCCT 536
Oy 37 ValValIleLeuProCysGlnHisAsnLeuCysArgLysCysAlaAsnAspValPheGln 56
Db 535 CGTGA---CTGCCATGCTCTCATCATTTTGTAGAATGTTTGAAGAAACATTCCTCAG 479
Oy 57 AlaSerAsnPro-----LeuTyrGlnSerArgGlySerThrIleValSerSerGly 74
Db 478 GCATCTGTAACCTTTATATATGAGACCTTTAGCAATTCGA----- 437
Oy 75 ArgPheArgCysProSerCysArgHisGluValIleuAspArgHisGlyValTyrGly 94
Db 436 ---CTCAAGTGCCCTTAATTTGCAGAGATATTTACTGAATTCCTCCAACTGGCATTCCT 380
Oy 95 LeuGlnArgAsnLeuLeuValGluAsnIleIleAspIleTyrLysGlnGluSerSerArg 114

```

```

Db 379 TTACCTGTTAATTTTGCATAAGGCTATATTGAAAGTACACAGCAAGAACCATCCA 320
Oy 115 ProLeuHisAlaLysAlaGluGlnHisLeuMetCysGluGlnHisGlnAspGlnLysIle 134
Db 319 GATATT-----GTCACTGCGCTTCAATTCAGGCAACCATTA 281
Oy 135 AsnIleTyrCysLeuSerCysGluValProThrCysSerLeuCysLysValPheGlyAla 154
Db 280 AATGTTACTGCTGTTATAGATAAAATTTACTTTGTGTCATTCCTTACCATAGTCA 221
Oy 155 HisLysAspCysGluValAlaProLeuProThrIleTyrLysArgGlnLysSerGluLeu 174
Db 220 CATCATGCTGTCATCTATAGATGACCTTCAAGTGCCCTATTGAAAGAAAGACACTCCT 161
Oy 175 SerAspGlyIleAlaMetLeuValAlaGlyAsnAspArgValGlnAlaValIleThrGln 194
Db 160 CAAAACCTGCTTAACACGTTG-----ACTGACACACACTGACAGATCTTACCAT 110
Oy 195 MetGluGluValAlaCysGlnThrIleGluAspAsnSerArgArg-----GlnLys 210
Db 109 CTATTGAAAAGCTGAAGAGAAACAAATCTCATCTGTGAGAAATGATCCAAAGCCATTAAG 50
Oy 211 GlnLeuLeuAsnGlnArgPheGluThrLeuCysAlaValLeuGluGlu 226
Db 49 GAAGCTGTTCTCCAGTATTTTAAGCAGCTTAATGATACATTAGAAGCAG 2

RESULT 12
US-09-764-864-1595/c
; Sequence 1595, Application US/09764864
; Patent No. US20020132753A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT223
; CURRENT APPLICATION NUMBER: US/09/764,864
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1792
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1595
; LENGTH: 599
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-864-1595

Alignment Scores:
Pred. No.: 5 23e-17 Length: 599
Score: 225.00 Matches: 59
Percent Similarity: 43.52% Conservative: 35
Best Local Similarity: 27.31% Mismatches: 98
Query Match: 11.77% Indels: 24
DB: 10 Gaps: 6

US-09-908-988b-2 (1-366) x US-09-764-864-1595 (1-599)
Oy 17 MetAspAsnLeuGluGlnLeuIleCysProIleCysLeuGluMetPheSerLysPro 36
Db 595 ATGCACAAATTTTGAGAGAGCTTAAGTATGTTATGTTTGAAGATTCCT 536
Oy 37 ValValIleLeuProCysGlnHisAsnLeuCysArgLysCysAlaAsnAspValPheGln 56
Db 535 CGTGA---CTGCCATGCTCTCATCATTTTGTAGAATGTTTGAAGAAACATTCCTCAG 479
Oy 57 AlaSerAsnPro-----LeuTyrGlnSerArgGlySerThrIleValSerSerGly 74
Db 478 GCATCTGTAACCTTTATATATGAGACCTTTAGCAATTCGA----- 437
Oy 75 ArgPheArgCysProSerCysArgHisGluValIleuAspArgHisGlyValTyrGly 94
Db 436 ---CTCAAGTGCCCTTAATTTGCAGAGATATTTACTGAATTCCTCCAACTGGCATTCCT 380
Oy 95 LeuGlnArgAsnLeuLeuValGluAsnIleIleAspIleTyrLysGlnGluSerSerArg 114

```

```

Db 379 TTACCTGTTAATTTTGCACCTAGGCGCTATTATTGAAAGACGACGACAGACCATCCA 320
Qy 115 ProleuHISAlaLysAlaGluGlnHisLeuMetCysGluGlnHisGluLysPgluysIle 134
Db 319 GATATTT-----GTCACTGCCCTGGAACATTACAGGCAACCATTA 281
Qy 135 AsnIleTyrcysLeuSerCysGluValProthrCysSerLeuCysLysValPheGlyAla 154
Db 280 AATGTTACTGCTATATAGATAAATAATTAGTTGTGTCATTCGCTTACCATAGTCA 221
Qy 155 HisLysAspCysGluValAlaProleuProthrIleTyrcysArgGlnLysSerGluLeu 174
Db 220 CATATGCTCATCTATATAGATGACCTCAAAAGTGCCTATTGAAAGAAAAGACACTCCT 161
Qy 175 SeraspGlyIleAlaMetLeuValAlaGlnHisAspArgValGlnAlaValIleThrGln 194
Db 160 CAAAACGCTGTAACACTG-----ACTGACACACACTGACAGATCTTACCAT 110
Qy 195 MetGluGluValCysGlnThrIleGluAspAsnSerArgArg-----GlnLys 210
Db 109 CTTATTGAAAAGCTGAAGAACAAAATCTCTCTGAGAAAATGATCCAGGCGATTAG 50
Qy 211 GlnLeuLeuAsnGlnArgPheGluThrLeuCysAlaValLeuGluGlu 226
Db 49 GAACCTGTTCTCAGTATTTTAAGAGGCTTAATGATACATTAGAACAG 2

```

## RESULT 13

```

US-09-764-877-2549
; Sequence 2549, Application US/09764877
; Patent No. US20020147140A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC005
; CURRENT APPLICATION NUMBER: US/09/764,877
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 4031
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2549
; LENGTH: 599
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-877-2549

```

## Alignment Scores:

```

Pred. No.: 5,23e-17 Length: 599
Score: 225.00 Matches: 59
Percent Similarity: 43.52% Conservative: 35
Best Local Similarity: 27.31% Mismatches: 98
Query Match: 11.77% Indels: 24
DB: 10 Gaps: 6

```

US-09-908-988b-2 (1-366) x US-09-764-877-2549 (1-599)

```

Qy 17 MetAspAsnLeuGlnLysGlnLeuIleCysProIleCysLeuGluMetPheSerLysPro 36
Db 5 ATGCACAAATTTTGGAGAGAGTAACTGTGCCATATGTTATAGATTTTCAACATCT 64
Qy 37 ValValIleLeuProCysGlnHisAsnLeuCysArgLysCysAlaAsnAspValPheGln 56
Db 65 CGGTGA---CTGCCATCTCTCATACATTTTGTGGAATAATTGTTGGAAACATCTCTCAG 121
Qy 57 AlaSerAsnPro-----LeuTrpGlnSerArgLysSerThrValSerSerGlyGly 74
Db 122 GCATCTGTAACCTTTATATATGAGACCTTACGAATTCCA----- 163
Qy 75 ArgPheArgCysProSerCysArgHisGlnValValLeuAspArgHisGlyValTyrcly 94
Db 164 ----CTCAAGTGCCTTAATTTGACAGAGTATTACTGAATTTGCTCCAAAGCGCATTTGAATCT 220
Qy 95 LeuGlnArgAsnLeuLeuValGluAsnIleIleAspIleTyrclyGlnLysSerSerArg 114

```

```

Db 221 TTACCTGTTAATTTTGCACCTAGGCGCTATTATTGAAAGTACACACAGACCATCCA 280
Qy 115 ProleuHISAlaLysAlaGluGlnHisLeuMetCysGluGlnHisGluLysPgluysIle 134
Db 281 GATATTT-----GTCACTGCCCTGGAACATTACAGGCAACCATTA 319
Qy 135 AsnIleTyrcysLeuSerCysGluValProthrCysSerLeuCysLysValPheGlyAla 154
Db 320 AATGTTACTGCTATATAGATAAATAATTAGTTGTGTCATTCGCTTACCATAGTCA 379
Qy 155 HisLysAspCysGluValAlaProleuProthrIleTyrcysArgGlnLysSerGluLeu 174
Db 380 CATATGCTCATCTATATAGATGACCTTCAAAAGTGCCTATTGAAAGAAAAGACACTCCT 439
Qy 175 SeraspGlyIleAlaMetLeuValAlaGlnHisAspArgValGlnAlaValIleThrGln 194
Db 440 CAAAACGCTGTAACACTG-----ACTGACACACACTGACAGATCTTACCAT 490
Qy 195 MetGluGluValCysGlnThrIleGluAspAsnSerArgArg-----GlnLys 210
Db 491 CTTATTGAAAAGCTGAAGAACAAAATCTCTCTGAGAAAATGATCCAGGCGATTAG 550
Qy 211 GlnLeuLeuAsnGlnArgPheGluThrLeuCysAlaValLeuGluGlu 226
Db 551 GAACCTGTTCTCAGTATTTTAAGAGGCTTAATGATACATTAGAACAG 598

```

## RESULT 14

```

US-09-764-877-2550
; Sequence 2550, Application US/09764877
; Patent No. US20020147140A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC005
; CURRENT APPLICATION NUMBER: US/09/764,877
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 4031
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2550
; LENGTH: 599
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-877-2550

```

## Alignment Scores:

```

Pred. No.: 5,23e-17 Length: 599
Score: 225.00 Matches: 59
Percent Similarity: 43.52% Conservative: 35
Best Local Similarity: 27.31% Mismatches: 98
Query Match: 11.77% Indels: 24
DB: 10 Gaps: 6

```

US-09-908-988b-2 (1-366) x US-09-764-877-2550 (1-599)

```

Qy 17 MetAspAsnLeuGlnLysGlnLeuIleCysProIleCysLeuGluMetPheSerLysPro 36
Db 5 ATGCACAAATTTTGGAGAGAGTAACTGTGCCATATGTTATAGATTTTCAACATCT 64
Qy 37 ValValIleLeuProCysGlnHisAsnLeuCysArgLysCysAlaAsnAspValPheGln 56
Db 65 CGGTGA---CTGCCATCTCTCATACATTTTGTGGAATAATTGTTGGAAACATCTTCTCAG 121
Qy 57 AlaSerAsnPro-----LeuTrpGlnSerArgLysSerThrValSerSerGlyGly 74
Db 122 GCATCTGTAACCTTTATATATGAGACCTTACGAATTCCA----- 163
Qy 75 ArgPheArgCysProSerCysArgHisGlnValValLeuAspArgHisGlyValTyrcly 94
Db 164 ----CTCAAGTGCCTTAATTTGACAGAGTATTACTGAATTTGCTCCAAAGCGCATTTGAATCT 220
Qy 95 LeuGlnArgAsnLeuLeuValGluAsnIleIleAspIleTyrclyGlnLysSerSerArg 114

```



```

Db 248 TTACCTGTTAATTTTGGACATAAGGCTATTATTGAAGATACAGCAAGACATCCCA 307
Oy 115 ProeuHisAlaValagluInHisLeuMetCysGluInHisGluAspGluPyluSile 134
Db 308 GATATT-----GTACCTGCGCCCTGAAACATATTACGGCAACCATTA 346
Oy 135 AsnIleTyrcysLeuSerCysGluValProThrCysSerLeuCysLysValPheGlyAla 154
Db 347 AATGTTTACTGTCTATAGATAAAATATTAGTTTGTGTCATTTGCTTACCAATAGTCAA 406
Oy 155 HisLysAspCysGluValAlaProleuProThrIleTyrcysLysGluInHisSerGluLeu 174
Db 407 CATATGCGTCACTCTATAGATGACCTTCAAGTGCCTATTTGAAGAAAGACACTCCT 466
Oy 175 SerAspGlyIleAlaMetLeuValAlaGluAsnAspArgValGlnAlaValIleThrGln 194
Db 467 CAAAACAGCTGTGAACACTTG-----ACTGCACACACTGACAGATCTTACCCCAT 517
Oy 195 MetGluGluValCysGluInThrIleGluAspAsnSerArgArg-----GlnLys 210
Db 518 CTATTGTACAAAGCTGAAGACAAATAATCTCATCTCTGAGAAATGATCCAGGCCATTAG 577
Oy 211 GlnLeuLeuAsnGlnArgPheGluThrLeuCysAlaValLeuGluGlu 226
Db 578 GAACTGTCTCCAGTATTTTAAGAGCTTAATGATACATTAGAACAG 625

```

## RESULT 17

US-09-728-952-1

```

; Sequence 1, Application US/09728952
; Patent No. US20020111302A1

```

## GENERAL INFORMATION:

```

; APPLICANT: Tang, Y. Tom
; APPLICANT: Zhou, Ping
; APPLICANT: Goodrich, Ryle
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Yamazaki, Yieki
; APPLICANT: Ujwal, Manusha L.
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. US20020111302A1el Nucleic Acids and
; FILE REFERENCE: 799
; CURRENT APPLICATION NUMBER: US/09/728,952
; CURRENT FILING DATE: 2000-11-30
; NUMBER OF SEQ. ID NOS: 101
; SOFTWARE: PC_FL_genes Version 2.0
; SEQ. ID NO 1
; LENGTH: 7596
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(7596)
; US-09-728-952-1

```

## Alignment Scores:

```

Pred. No.: 3, 65e-15 Length: 7596
Score: 223.50 Matches: 85
Percent Similarity: 39.07% Conservative: 67
Best Local Similarity: 21.85% Mismatches: 165
Query Match: 11.69% Indels: 72
DB: 10 Gaps: 13

```

US-09-908-988b-2 (1-366) x US-09-728-952-1 (1-7596)

```

Oy 14 AlaHisAsnMetAspAsnLeuGluLysGluInLeuIleCys--ProIleCys--LeuGluMet 32
Db 2805 GCTCAT-----CATCTCCCGCGCGCAAAATCTGCTCTGCTGCTCCCTTAGTTT 2855
Oy 32 tPheserLysProValValIleLeuProCysGlnHisAsnLeuCysArgLysCysAlaAs 52

```

```

Db 2856 CACCATGCAGGACAGTCAATA-----TGCTCTCA 2882
Oy 52 nAspValPheGlnAlaSerAsnProLeuTPrGlnSerArgLysSerThrValSerSe 72
Db 2883 AAACTATATCCCTCCAG-----ACCTGACGCTATCC-- 2916
Oy 72 rGlyIArgPheArgProSerCysArgHisGluValIleLeuAspArgHisGlyVa 92
Db 2917 -----TGTCAGTATGCGGCGACAGCTCCATCTCCACAGACAGGGGCT 2960
Oy 92 lTyrcylLeuGlnIArgAsnLeuValGluAsnIleLeaspIleTyrcylLysGln----- 110
Db 2961 CTCGCGACTGCGAACAACCTTCTTCATCAGCAGCCTCATGGAGGACATGACAGCAGCACC 3020
Oy 111 -----GluSerSerArgProLeuHisAlaLysAlaGluInHisle 124
Db 3021 TGATGGGGCCACGACCCCGGAGACCCGCCCTCAGTGTAGTGGCTGGCCCTCT 3080
Oy 124 uMetCysGluGluInHisGluAspGluLysIleAsnIleTyrcysLeuSerCysGluValPr 144
Db 3081 CTCCTGCCCCAACCATGATAGGCAAGACGATGAGTTTACTGTGAGGCTGTGAGACGCG 3140
Oy 144 oThrCysSerLeuCysLysValPheGlyAlaHisLysAspCysGluValAlaProLeuPr 164
Db 3141 CATGTGTGTAGTGGCGCGCC---GGGAGACATGTGAGCATGGCACAGTGTCTGAG 3197
Oy 164 oThrIleTyrcysArgGlnLysSerGluLeuSerAspGlyIleAlaMetLeuValAlaGl 184
Db 3198 GATGTGTGTAGTGGCGCGCCCTGACAGCCCTGACAGCTGAGTGTGGCGCG 3257
Oy 184 yAsnAspArgValGlnAlaValIleThrGlnMetGluGluValCysGlnThrIleGluAs 204
Db 3258 ATTGCCACAGCTGTCCGACAGCAATTCCTTGTGTGGGGCATCCAGCCAGCTGCAGGA 3317
Oy 204 pAsnSerArgArgGlnLysGluLeuLeuAsnGlnArgPheGluThrLeuCysAlaValIle 224
Db 3318 GCCCAAGCGCAGAGCCCTGCGCCAGATCATGTCAGCGCTTCAGAGACCTGGAGCAAGCACT 3377
Oy 224 uGluGluArgLysGlyLysGluLeu-----LeuGlnAlaLeuAlaArgGluGlnGlu 241
Db 3378 GCAGCAGCGCAGAGCCTGTGTGTCAGCAGCAGCTGTGAGCCATTTGTGGGGCAACAGAA 3437
Oy 241 uLysLeuGlnArgValArgLysLeuIleArgGlnTyrcylAspHisLeuGluGlySerSe 261
Db 3438 GGTGTTCAAAGCCAGCTGTGACACACTGTGCGCAGGCTGAGGACATCGCAGTAGCTG 3497
Oy 261 rLysLeuValGluSerAlaIleGlnSerMetGluGluProGlnMetAlaLeuTyrcyl 281
Db 3498 CAGCTTTCAGAGCAGCAGCAGCTGCGCTGGGCTCGGCCCGAGGCTGTGTCGGCGCA 3557
Oy 281 nGlnAlaLysGluLeuIleAsnLysValGlyAlaMetSerLysValGluLeuAlaGlyAr 301
Db 3558 GCACATGCGAGAG-----CGGCTGGCTGATTTGGCGGCACAGGCCCTTCCGGAGCG 3608
Oy 301 qProGluProGlyTyrcylu-----SerMetGluGlnPheSerValSerVa 316
Db 3609 GCCACATGAGATATGCACACTGTGACTGTCTTGTAGGTGACGCTGTGCGGCATCGT 3668
Oy 316 lGluHisValAlaGluMetLeuArgThr----- 325
Db 3669 GCTCAATCTGTGGCGACTGCTCACACAGGCGCAGTGCACAGCAACAGGTGGCCAGCGG 3728
Oy 326 -----lLeasPheGlnProGlyAlaAlaGluLysArgGluAspAs 339
Db 3729 AGAGGGCTGTGGCCAGGCGCTAGTGGGCGCAGCTGCTGTCACTGTCACTACCAAGAA 3788
Oy 339 pAspMetAlaLeuAspGlyLysGluGluGlyAsnAlaGlyLeuGluGluGluArgLysAspVa 359
Db 3789 CAGG-----GACGGGGGGTGTGTGTGGCCACAGGACGCGTCAAGCTGTGACAGAT 3839
Oy 359 l-----ProGluGlySerGlyLeu 365

```









GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OW protein - nucleic search, using frame\_plus\_p2n model

Run on: December 3, 2002, 12:14:13 ; Search time 56 Seconds  
(without alignments)  
2004.352 Million cell updates/sec

Title: US-09-908-988B-2  
Perfect score: 1912  
Sequence: 1 MFTVGFKPLGDHNMNDL.....EGNAGLEERLDVPGSLH 366

Scoring table:  
BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 441362 seqs, 153338381 residues  
Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=framet-p2n.model -DEV=xih  
-O=/cgnt2.1/USPTO.spool/US09908888/runat\_26112002\_112154\_11213/app.query.fasta\_1.519  
-DB=Issued\_Patents\_NA -OPMT=fastat -SUFFIX=rni -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOCALLIGN=200 -THR.SCORE=pct -THR.MAX=100 -THR.MIN=0 -ALIGN=20  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09908888 @CGN 1.1.17 @unat\_26112002\_112154\_11213 -NCPU=6 -ICPU=3  
-NO\_XLPPY -NO\_MMP -IAREQUERY -NEG\_SCORES=0 -WAIT -LONGLOG -DEV.TIMEOUT=7  
-WARN.TIMEOUT=30 -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6 -Fgapext=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database: Issued\_Patents\_NA.\*  
1: /cgnt2\_6/ptodata/1/ina/5A.COMB.seq.\*  
2: /cgnt2\_6/ptodata/1/ina/5B.COMB.seq.\*  
3: /cgnt2\_6/ptodata/1/ina/6A.COMB.seq.\*  
4: /cgnt2\_6/ptodata/1/ina/6B.COMB.seq.\*  
5: /cgnt2\_6/ptodata/1/ina/PCNUS.COMB.seq.\*  
6: /cgnt2\_6/ptodata/1/ina/backfile1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	1101.5	57.6	1757	4	US-09-484-970B-37
2	255.5	13.4	3262	4	US-09-561-989-9
3	201	10.5	2854	4	US-08-724-394A-17
4	178	9.3	3446	4	US-09-484-970B-93
5	166.5	8.7	3312	1	US-08-049-473-1
6	166.5	8.7	3312	1	US-08-312-648-1
7	166.5	8.7	3312	5	PCT-US94-04190-1
8	132	6.9	2155	2	US-08-095-728B-5
9	132	6.9	2155	5	PCT-US97-02320A-5
10	132	6.9	3511	3	US-08-892-747-13
11	131	6.9	246240	2	US-08-724-394A-20
12	131	6.9	246240	2	US-08-724-394A-21

13	131	6.9	246240	2	US-08-724-394A-22	Sequence 22, Appl
14	127	6.6	1767	1	US-07-903-466-2	Sequence 2, Appl
15	127	6.6	1767	5	PCT-US93-05794-2	Sequence 2, Appl
16	127	6.6	3018	1	US-07-903-466-1	Sequence 1, Appl
17	127	6.6	3018	5	PCT-US93-05794-1	Sequence 1, Appl
18	123	6.4	5173	4	US-08-801-308-2	Sequence 2, Appl
19	122.5	6.4	2240	4	US-08-697-610-1	Sequence 1, Appl
20	122.5	6.4	2240	4	US-08-349-357-1	Sequence 1, Appl
21	122.5	6.4	2301	4	US-09-085-199B-8	Sequence 8, Appl
22	122	6.4	3036	2	US-08-306-691B-52	Sequence 52, Appl
23	122	6.4	3036	5	PCT-US92-02320A-1	Sequence 1, Appl
24	122	6.4	3036	5	PCT-US92-02320A-1	Sequence 1, Appl
25	118.5	5.9	2262	4	US-08-639-237-1	Sequence 2, Appl
26	113.5	5.9	2248	1	US-08-975-405-1	Sequence 1, Appl
27	113.5	5.9	2264	4	US-09-167-109-6	Sequence 6, Appl
28	113.5	5.9	2121	2	US-08-897-340-6	Sequence 6, Appl
29	112.5	5.9	2121	3	US-09-252-329-6	Sequence 6, Appl
30	112.5	5.9	2121	5	PCT-US95-16980-1	Sequence 1, Appl
31	112.5	5.9	2359	5	US-08-074-879-9	Sequence 9, Appl
32	112	5.9	7833	1	US-08-468-057A-9	Sequence 9, Appl
33	112	5.9	7833	1	US-08-468-057A-9	Sequence 9, Appl
34	111.5	5.8	2455	4	US-09-167-109-3	Sequence 3, Appl
35	111	5.8	4722	4	US-08-979-608A-14	Sequence 14, Appl
36	111	5.8	5707	1	US-08-425-061-11	Sequence 11, Appl
37	111	5.8	5707	1	US-08-825-886-11	Sequence 11, Appl
38	111	5.8	5709	1	US-08-425-061-7	Sequence 7, Appl
39	111	5.8	5709	1	US-08-425-061-8	Sequence 8, Appl
40	111	5.8	5709	1	US-08-425-061-9	Sequence 9, Appl
41	111	5.8	5709	1	US-08-825-886-7	Sequence 7, Appl
42	111	5.8	5709	1	US-08-825-886-8	Sequence 8, Appl
43	111	5.8	5709	1	US-08-825-886-9	Sequence 9, Appl
44	111	5.8	5710	1	US-08-425-061-6	Sequence 6, Appl
45	111	5.8	5710	1	US-08-825-886-6	Sequence 6, Appl

## ALIGNMENTS

RESULT 1  
US-09-484-970B-37  
; Sequence 37, Application US/09484970B  
; Patent No. 6426186  
; GENERAL INFORMATION:  
; APPLICANT: Jones, Karen A.  
; APPLICANT: Volkuth, Wayne  
; APPLICANT: Walker, Michael G.  
; TITLE OF INVENTION: BONE REMODELING GENES  
; FILE REFERENCE: PB-0014 US  
; CURRENT APPLICATION NUMBER: US/09/484,970B  
; CURRENT FILING DATE: 2000-01-18  
; NUMBER OF SEQ ID NOS: 172  
; SOFTWARE: PERL Program  
; SEQ ID NO 37  
; LENGTH: 1757  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc.feature  
; OTHER INFORMATION: Incyte ID No. 6426186 247384.1CB1  
US-09-484-970B-37

## Alignment Scores:

Pred. No.: 4.9e-118  
Score: 1101.50  
Percent Similarity: 78.29%  
Best Local Similarity: 60.86%  
Query Match: 57.61%  
DB: 4  
Gaps: 4

US-09-908-988B-2 (1-366) x US-09-484-970B-37 (1-1757)

Qy 10 LeuLeuGlyAspAlaHisAsnMetAspAsnLeuGluysGlnLeuIleGysProIleGys 29  
Db 143 CTGATCCAGATGGGATCCATCCATGAGAGAACTTGAGAGAAAGCAGCTGATCTGCTATCTGCC 202

30 LeuGlumetPheSerLysProValValIleuProCysGlnHisAsnLeuCysArgLys 49  
 203 CTGAGATGTTTACCAAGCCAGTGTATCTTCCGTCGACACACACTGTCCCGAAG 262  
 50 CysAlaAsnAspValPheGlnAlaSerAsnProLeuTyrPglSerArgLysThr 69  
 263 TGTGCCAATATATTTTCCAGGCTCTTAACCCGTATTTCGCCCAAGAGAGTACCACC 322  
 70 ValSerSerGlyArgPheArgCysProSerCysArgHisGlnValValLeuAspArg 89  
 323 ATGGCATCAGGGGCGGATTCGGCTGGCCATCTGTAGACATGATGGTTTGGATTAGA 382  
 90 HisGlyValTyrGlyLeuGlnArgAsnLeuValGluAsnIleLeuAspIleTyrLys 109  
 383 CATGGGTATATGACTTCAAGAGAACTGCTGGTGGAGAACATCATCTGCATCTCAAAA 442  
 110 GlnGlu--SerSerArgProLeuHisAlaLysAlaGlnGlnHisLeuMetCysGlu 128  
 443 CAGAGTGTCTCCAGTCCGCGCTG-----CAGAAGGGCAGTCCACCCATGTGCAAGGAG 496  
 129 HisGlnAspGlnLysIleAsnIleTyrCysLeuSerCysGlnValProThrCysSerLeu 148  
 497 CACGAATGAGAAATCAATCACTACTGTCTCAGCTGAGGTGCGCCACCTGCTCCATG 556  
 149 CysLysValPheGlyAlaHisLysAspCysGlnValAlaProLeuProThrIleTyrLys 168  
 557 TGCAGAGTGTGGGATTCACACAGGCTGCGAGGTGCCCCCATGTGCAGATGCTTCCAG 616  
 169 ArgGlnLysSerGlnLeuSerAspGlyIleAlaMetLeuValAlaGlnAsnAspArgVal 188  
 617 GGCAGAAAGACTGAACATAAATGATCTCTCCATGCTGCTGGCGGGAATGACCGGTG 676  
 189 GlnAlaValIleThrGlnMetGlnGlnValCysGlnThrIleGlnAspAsnSerArgArg 208  
 677 CAGACCATCATCTACTGAGGATTCCTCGAGTACCAAGGAGGACAGTACACAG 736  
 209 GlnLysLeuLeuAsnGlnArgPheGlnThrLeuCysAlaValLeuGlnArgLys 228  
 737 GTAAAGAGAGAGCTGAGCAGACAGATTGTGACGTTGATGCCATCTCGATGAGAAAGAA 796  
 229 GlyGlnLeuLeuGlnAlaLeuAlaArgGlnGlnGlnLysLeuGlnArgValArgGly 248  
 797 AGTGAAGTGTGCTGACGCGATCCGACGAGGAGGAGAAAGCTTCACTTCGAGGACC 856  
 249 LeuIleArgGlnTyrGlyAspHisLeuGlnGlySerSerLysLeuValGlnSerAlaIle 268  
 857 CTCATCCAGCAGTACACAGAGACGTGGACAAAGTCCACAAAGCTGGTGAAGTCCATC 916  
 269 GlnSerMetGlnGlnProGlnMetAlaLeuTyrLeuGlnGlnAlaLysGlnLeuIleAsn 288  
 917 CAGTCCCTGAGCAGACCTGGGGAGCCACCTCTCTGACTCCCAAGCAATCATCAAA 976  
 289 LysValAlaGlyAlaMetSerLysValGlnLeuAlaGlyArgProGlnProGlyTyrGlnSer 308  
 977 AGCATTTGGAGACTTCCAGGGCTGCCAGCTGGGGAGACACAGAGCGGCTTGAAGAAC 1036  
 309 MetGlnGlnPheSerValSerValGlnHisValAlaGlnMetLeuArgThrIleAspPhe 328  
 1037 ATGCACTTCTTACTTGTGATTAGACACATAGCAGAGCCCTGAGAGCATGACTTTT 1096  
 329 GlnProGlyAlaAlaGlyAspGlnLeuAspAspMetAla-----LeuAsp 344  
 1097 -----GGAGACAGATGAGAGAGAGAGAAATTCATTGAAGAAAGACATCAGGAA 1144  
 345 GlyGlnGlnGlnGlnAlaGlyLeuGlnGlu 354  
 1145 GAGGAGAGTCCACAGAAAGGAGAAAGAA 1174

RESULT 2  
 US-09-561-989-9  
 ; Sequence 9, Application US/09561989  
 ; Patent No. 6468750

; GENERAL INFORMATION:  
 ; APPLICANT: KOLLER, Klaus-Peter  
 ; TITLE OF INVENTION: No. 6468750e1 Cell Regulation Factor T1020  
 ; FILE REFERENCE: seq  
 ; CURRENT APPLICATION NUMBER: US/09/561, 989  
 ; CURRENT FILING DATE: 2000-04-27  
 ; NUMBER OF SEQ ID NOS: 10  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 9  
 ; LENGTH: 3262  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; US-09-561-989-9  
 Alignment Scores:  
 Pred. No.: 2,46e-19 Length: 3262  
 Score: 255.50 Matches: 67  
 Percent Similarity: 41.73% Conservative: 44  
 Best Local Similarity: 25.19% Mismatches: 102  
 Query Match: 13.36% Indels: 53  
 DB: Gaps: 2  
 US-09-908-988b-2 (1-366) x US-09-561-989-9 (1-3262)  
 78 CysProSerCysArgHisGlnValValLeuAspArgHisGlyValTyrGlyLeuGlnArg 97  
 37 TGCCTCCATGACACCGCAGCCATCTGATGACCGGGGGCTCCGCTCCCAAG 96  
 98 AsnLeuValGlnAsnIleLeuAspIleTyrGlnGlnIleuSerArgProLeu--- 116  
 97 AATCGGATGTGAAGGGGTAATTGACCGCTACCAAGAGAGCAAAAGCGCGCCCTCAAG 156  
 117 -----HisAlaLysAlaGlnGlnHisLeuMetCysGlu----- 127  
 157 TCCCAAGCTCGGAGAGAGCGCCCAAGAGAGCCAGCTCATGTGCGAAGTGGATGTC 216  
 127 ----- 127  
 217 TTTACTGCGATCCGTGCGGCTGCGCTGCCACCGCCCGGGGGCCCTAGCCAGACAG 276  
 127 ----- 127  
 277 CCGCTGCTGCCCCCGCCAGGCTCGTGTACCGCGAGCTGAGCCAGCAAGTCTCC 336  
 128 -----GlnHisGlnAspGlnLysIleAsnIleTyrCysLeuSerCysGlnValPro 144  
 337 ACGTCACAGACACGAGCGTGAAGAACCAACAGCATGTACTCGTCAATGCAAGATGCC 396  
 145 ThrCysSerLeuCysLysValPheGlyAlaHisLysAspCysGlnValAlaProLeuPro 164  
 397 GTGTCTTACCAAGTGTGGAGAGGGGCAACACTCCAGCCAGCAAGTCAAGCTGTGGGG 456  
 165 ThrIleTyrLysArgGlnLysSerGlnLeuSerAspGlyIleAlaMetLeuValAlaGly 184  
 457 GCCATGTGAACATACATAGACGACCTCTCCAGGCGGTGAAGGAGCTGCACAGAG 516  
 185 AsnAspArgValGlnAlaValIleThrGlnMetCysGlnValCysGlnThrIleGlnAsp 204  
 517 GCCAAAGAACCCAGAGAGTTCGTGTGACAGTCCGCAACATGTGTCACACAGATCCAGAG 576  
 205 AsnSerArgArgGlnLysGlnLeuLeuAsnGlnArgPheGlnThrIleCysAlaValLeu 224  
 577 AACAGTGTGAGATTGAAGCTGTGTGTGGGCCCAATGTGATGCCCTATGATGCCCTC 636  
 225 GlnGlnArgGlyGlnGlnLeuLeuGlnAlaLeuAlaArgGlnGlnGlnLysLeuGln 244  
 637 AACAGAAAGAAACCCAGCGTGTGGCCCGGTCAACAAGAGAGCATACACAGCTGAAG 696  
 245 ArgValArgGlyLeuIleArgGlnTyrGlyAspHisLeuGlnGlySerSerLysLeuVal 264  
 697 GTGTTCGAGATCAGATCTCTCACTGCACAGTGAATTTGCCACAGACCAAGGTCTCATG 756  
 265 GlnSerAlaIleGlnSerMetGlnGlnProGlnMetAlaLeuTyrLeuGlnGlnAlaLys 284

DB 757 GAGTACTGCTGGAGGTGATTAGGAAATGATCCTAGTGTGTTTTCAGATTCTGCAC 816  
OY 285 GtuleuIleasnLysVal 290  
DB 817 GCCCTCATAGAAAGAGTG 834

RESULT 3  
US-08-724-394A-17  
Sequence 17, Application US/08724394A  
Patent No. 5872237  
GENERAL INFORMATION:  
APPLICANT: Feder, John N.  
APPLICANT: Krommal, Gregory S.  
APPLICANT: Launer, Peter M.  
APPLICANT: Ruddy, David A.  
APPLICANT: Thomas, Winston  
APPLICANT: Tsuchihashi, Zenta  
APPLICANT: Wolf, Roger K.  
TITLE OF INVENTION: Megabase Transcript Map: No. 5872237el  
TITLE OF INVENTION: Sequences and Antibodies Thereo  
NUMBER OF SEQUENCES: 31  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: TOWNSEND and TOWNSEND and CREW LLP  
STREET: Two Embarcadero Center, 8th Floor  
CITY: San Francisco  
STATE: CA  
COUNTRY: USA  
ZIP: 94111-3834  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/724,394A  
FILING DATE: 01-Oct-1996  
CLASSIFICATION: 536  
ATTORNEY/AGENT INFORMATION:  
NAME: Fitts, Renee A.  
REGISTRATION NUMBER: 35,136  
REFERENCE/DOCKET NUMBER: 017957-000100  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-576-0200  
TELEFAX: 415-576-0300  
INFORMATION FOR SEQ. ID NO: 17:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2854 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: not relevant  
TOPOLOGY: not relevant  
MOLECULE TYPE: cDNA  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: 1..2854  
OTHER INFORMATION: /note= "cDNA 27"  
US-08-724-394A-17

Alignment Scores:  
Pred. No.: 4.25e-13 Length: 2854  
Score: 201.00 Matches: 90  
Percent Similarity: 38.12% Conservative: 56  
Best Local Similarity: 23.50% Mismatches: 161  
Query Match: 10.51% Indels: 76  
DB: 2 Gaps: 14

US-09-908-988b-2 (1-366) x US-08-724-394A-17 (1-2854)  
OY 22 LysGlnIleuIleCysProIleCysLeuGluMetPheSerIysProValIleLeuPro 41  
DB 55 GAGGAAGCCACGCTGCTCATCTGCTGAGCTGATGAGCAACCCAGTAGAGATC--AAC 111  
OY 42 CysGlnHisAsnLeuCyArGlyCysAlaAsnAspValPheGlnAlaSerAsnProLeu 61

DB 112 TGTGACACAGCTACTGCTCCACTGTGTATACAGACTTCTTTAA-----AACCCAGC 165  
OY 62 TrpGlnSerArgGlySerThrThrValSerSerClyGlyArgPheArgCysProSerCys 81  
DB 166 CAAGAAGCAACTAGCGCAGAGACA-----TTCTGCTGTCCCGCAGTGT 207  
OY 82 ArgHisGluValIleuAspArgHisGlyValIleuArgGlyLeuGlnArgAsnLeuVal 101  
DB 208 CGGGCTCCATTTCAATATGAT-----AGCTCCGACCCACACAGCAGCTG 252  
OY 102 GluAsnIleIleAspIleTyrIysGlnGluSerSerArgProLeuHisAlaLysAlaGlu 121  
DB 253 GGAAGCCTCATTTGAAGCCCTCAAGAG-----ACGAT 285  
OY 122 GlnHisLeuMetCysGlnGlnGlnHisGluAspGluLysIleAsnIleTyrCysLeuSerCys 141  
DB 286 CAGAAATGTCAATGATGAGAACAC--GGAGAGCAGTTCACCTGTTCTCGAAGACGAG 342  
OY 142 GluValProThrCysSerLeuCysLysValPheGlyAlaHisLysAspCysGluValAla 161  
DB 343 GGGCAGCTCATCTGCTGGGGCTGTGAGCGGCACACAGCACAAGGCGCACACACAGCT 402  
OY 162 ProLeuProThrIleTyrIysArgGlnLysSerGluLeuSerAspGlyIleAlaMetLeu 181  
DB 403 CTGTGTGACAGAGTATGCCAGCGCTACAGGAAGAAAGCTCAGAAAGCTGTGACAAACTG 462  
OY 182 ValAlaGlyAsnAspArg-----ValGlnAlaValIleThrGln 194  
DB 463 AAGCAACTTGAAAGACAGATGTACGAGCAGAGACTGTCCACACCAATGCAGATTAACATAA 522  
OY 195 MetGluGluValCysGlnThrIleGluAspAsnSerArgArgGlnLysGlnLeuAsn 214  
DB 523 TGGAAAGAAAGGTACAGATT-----CAGACCAAAAAATCCGG 561  
OY 215 GlnArgPheGlnThrLeuCysAlaValLeuGluGluArgLysGlyGluLeuGlnAla 234  
DB 562 TCTGACTTTAAGAAATCTCCAGTGTTCCTACATGAGAGAAAGATCTTACTCTGGAG 621  
OY 235 LeuAlaArgGluGlnGluGlnLysLeuGlnArgValArgLysLeuIleuGlnIleTyrGly 254  
DB 622 CTGAGAGAAAGAACACACAGACTGTGAGTGA-----CTGAGGAGACTATGAG 669  
OY 255 AspHisLeuGluGlnLysSerSerLysLeuValGluSerAlaIleGlnSerMetGluGluPro 274  
DB 670 GCTGCTCTGGGCGTGAAGGATGAAGCACTGAAGTCAAGACCCACTCTCGGAACCGAGAAAA 729  
OY 275 -----GlnMetAlaLeuTyrLeuGlnGlnAlaLysGluLeuIleAsnLysValGly 291  
DB 730 TGTCAAGGCTCAGCCCAAGAAATTGCTGCAGAAATGCAATGACACTTGTGACAGAGTTGG 789  
OY 292 AlaMetSerLysValGluLeuAlaGlyArgProGluProGlyIleGluSerMetGluGln 311  
DB 790 GCTGTG--AAGCTGGAACATCAGAGGCTGTCTCTGGAATCTCATATGTGCAAT 846  
OY 312 PheSerValSerValGlnHisValAlaGluMetLeuArg-----ThrIleAsp 327  
DB 847 GTTTCACAGCTTACTTCGATGATGAGAAATGTTAAGAGTCAATCAAGTTAGTGTACT 906  
OY 328 PheGlnProGlyAlaAlaGlyAspGlu-----GluAspAspMetAla 336  
DB 907 CTGGATCCAGATACAGCTCTCAGCAACTAATTCTCTGAGATCGAGACAGATGACT 966  
OY 337 -----GluAspAspMetAla----- 342  
DB 967 CGTGAATACACCCAGGAGATTCAGGACACATCTTCCAGGAGATTTACTGCCCTCCCTGT 1026  
OY 343 --LeuAspGlyGluGlnGlyAsnAlaGlyLeuGluGlnGluArgLeuAspValProGlu 361  
DB 1027 GCTTGGGTGTGAAAGGCTTCACCTCAGGAAGAGCTTGAAGAGGATGTTGGGAA 1086  
OY 362 GlySerGly 364













Percent Similarity: 33.17% Conservative: 51  
 Best Local Similarity: 20.91% Mismatches: 154  
 Query Match: 6.90% Indels: 124  
 DB: 5 Gaps: 17

US-09-908-988b-2 (1-366) x PCT-US92-02320A-5 (1-2155)

QY 24 LeuileCysProIleCysLeuGluMetPheSerLysProValAlaIleuProCysGln 43  
 DB 243 CTGGCGCTCCAGCAATGC---CAGCGCGAAGCCCAAGTCCCAAGCTGCTGCTGCTG 299  
 QY 44 HisAsnLeuCysArgLysCysAlaAsnAspValPheGlnAlaSerAsnProLeuTyrGln 63  
 DB 300 CACACGCTGTGCTGAGATGCTGAG---GAGTCCGCGGCGGCGGCGGCGGCGGCGGCG 326  
 QY 64 SerArgLysSerThrThrValSerSerGlyArgPheArgCysProSerCysArg--- 82  
 DB 327 -----GCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 359  
 QY 83 -----HisGluValAlaLeuAspArgHisGlyValTyrGlyLeuGln 96  
 DB 360 CCGTGGCCCTAGTGCAGACACACCCGCTGATACGCTTTTTCGAGAGTGTGAG 419  
 QY 97 ArgAsnLeuValAlaGlnAsnIleIle----- 105  
 DB 420 CGGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 479  
 QY 106 -----AspIleTyrLysGlnLysSerArgProLeuHisAlaLysAla---Glu 121  
 DB 480 GAGTGCAGCGACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 539  
 QY 122 GlnHisLeuMetCysGlnGlnHisGlnAspGlyLysIle----- 134  
 DB 540 GCACACACAGTGTCTTCAAGCAGAGGCGCGCGCTGACAGAGCTGCGCAACAGTGC 599  
 QY 134 ----- 134  
 DB 600 GTGGGTGAGTCTCGAGAGCGACCCCGCAACACATCTTCTGCTCCAAACCCCAAC 659  
 QY 135 -----AsnIleTyrCysLeuSerCysGlnValProThrCysSer 147  
 DB 660 CACCGCACCCCTACGCTGACACGATCTACTGCGAGATGTTTCAAGCGGCTGCTGCTGCTG 719  
 QY 148 LeuCysLysValPhe---GlyAlaHisLysAsp---CysGlnValAlaProLeuPro 164  
 DB 720 TCGTGGCGGCTCTTGAACAGCAGCAGTGAAGTCAAGTGGACATCAGC-----GCA 773  
 QY 165 ThrIleTyrLysArgLysGlnLysSerGluLeuSerAspGlyIleAlaMetLeuValAlaGly 184  
 DB 774 GAGATCCAGCAGCGACAGAG---GAGGTGACCGCAGCAGCAGCGGCTGCGAGAGCAG 830  
 QY 185 AsnAspArgValAlaGlnAlaValIleThrGlnMetGluGlnValCysGlnThrIleGluAsp 204  
 DB 831 GATAGTGCCTTTGGCGGCTTCAAGCAGATGACACGCGCGCTGCGCAAGTGGGCGCG 890  
 QY 205 AsnSerArgArgLysGlnLysGlnLeuAsnGlnArgPheGluThrLeuCysAlaValLeu 224  
 DB 891 GCGCGTCCAGACAGCGGAGCGAGCTGATCCGCGAGCGGCTGCGCAGAGTGAAGTCAAGCG 950  
 QY 225 GlnGluArgLysGlnLysGlnLeuGlnAlaLeu-----AlaArgLysGln 240  
 DB 951 CGGCGTCCAGAGCGGAGCGAGCTGCTGAGAGCTGTGAGCGCGGCTGACACGCGACTACGAG 1010  
 QY 241 GlnLys-----LeuGlnArgValArg----- 247  
 DB 1011 GAGATGAGCAGTGGCTGGCGCGCTGATGCTGCTGCTGAGCGCATCCGACGCGAGC 1070  
 QY 248 -----Gly 248  
 DB 1071 GCGGTGTCAGAGAGATGATGCTAGCCTCGAGACAGAGGCTGCGACATGACACGCT 1130  
 QY 249 LeuIleArgGlnTyrGlnLysAspHisLeu---GlnGlySerSerLysLeuValGlnSerAla 267  
 DB ----- 267

DB 1131 TTCCTGGCCAGGAGGCGCTCTGCGCGCTGCGCCAGAGAGGAGCCCGACAGCTGCC 1190  
 QY 268 IleGlnSerMetGluGlnProGlnMetAlaLeuTyrLeuGlnAlaLysGlnLeuIle 287  
 DB 1191 GTGGCAGCCAGATGCTGCTGAGATGCTGAGATGCTGAGATGCTGAGATGCTGAGATGCT 1250  
 QY 288 AsnLysValAlaMetSerLysValGlnLeuAlaGlyArgProGlnProGlyTyrGln 307  
 DB 1251 ACCCAG---GGGAAAGATGACAGCTGTATCCAAAGAAAGCCAGCCAGAGGCTGCCAGC 1307  
 QY 308 SerMetGlnGlnPheSerValSerValGlnHisValAlaGlnMetLeuArgThrIleAsp 327  
 DB 1308 CCGAGGACCCCTATGAGCTGTGACCTGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1361  
 QY 328 PheGlnProGlyAlaAlaGlyAspGlnGlnAspAspAspMetAlaLeu 343  
 DB 1362 ---CAGTTTCAGGCGCTGCGCTGCTGAGAGCCAGCCATGCTGTG 1406

RESULT 10  
 US-08-892-747-13  
 ; Sequence 13, Application US/08892747  
 ; Patent No. 6057153  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Shaji T. George, Michael Ma, Martina Werner,  
 ; APPLICANT: Umberto Pace and Allan R. Goldberg  
 ; TITLE OF INVENTION: Stabilized External Guide Sequences  
 ; NUMBER OF SEQUENCES: 50  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Patrea L. Pabst  
 ; STREET: 2800 One Atlantic Center  
 ; CITY: Atlanta  
 ; STATE: Georgia  
 ; COUNTRY: USA  
 ; ZIP: 30309-3450  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patent Release #1.0, Version #1.25  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/892,747  
 ; FILING DATE:  
 ; CLASSIFICATION: 435  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: 08/372,556  
 ; FILING DATE: January 13, 1995  
 ; CLASSIFICATION: 435  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: PCT/US96/00513  
 ; FILING DATE: January 19, 1996  
 ; CLASSIFICATION: 435  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Pabst, Patrea L.  
 ; REGISTRATION NUMBER: 31,284  
 ; REFERENCE/DOCKET NUMBER: IL109CIP2  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (404) 873-8794  
 ; TELEFAX: (404) 873-8795  
 ; INFORMATION FOR SEQ ID NO: 13:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 3511 base pairs  
 ; TYPE: nucleic acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: cDNA  
 ; HYPOTHETICAL: NO  
 ; ANTI-SENSE: NO  
 ; FEATURE:  
 ; NAME/KEY: misc\_feature  
 ; LOCATION: 1..3511  
 ; OTHER INFORMATION: /function- "PML-RAR DNA Sequence."  
 ; US-08-892-747-13



## Alignment Scores:

Pred. No.:	0.0753	Length:	246240
Score:	131.00	Matches:	37
Percent Similarity:	42.96%	Conservative:	21
Best Local Similarity:	27.41%	Mismatches:	53
Query Match:	6.85%	Indels:	24
DB:	2	Gaps:	6

US-09-908-988b-2 (1-366) x US-08-724-394A-20 (1-246240)

QY 22 LysGlnLeuIleCysProIleCysLeuGluMetPheSerLysProValIleLeuPro 41

Db 71769 GAGGAAGCACCCTGCTCCATTCGCTGAGCCCTGATGACCAACCAAGCAATC---AAC 71825

QY 42 CysGlnHisAsnLeuCysArgLysCysAlaAsnAspValPheGlnIleAsnProLeu 61

Db 71826 TGTGACACAGCTACTGCTGCTGATATACAGACTTCTTAA-----AACCCAGC 71879

QY 62 TrpGlnSerArgLysThrValSerSerGlyArgPheArgCysProSerCys 81

Db 71880 CAAAGCAACGACGACGACGAC-----TTCTGCTGTCTCCAGCTGT 71921

QY 82 ArgHisGluValIleuAspArgHisGlyValTyrGlyLeuGlnArgAsnLeuVal 101

Db 71922 CGGCTCCATTCATATATGAT-----AGCCTCGACCAACAAGCAGCTG 71966

QY 102 GluAsnIleIleAspIleTyrLysGlnGluSerSerArgProLeuHisAlaLysAlaGlu 121

Db 71967 GGAAGCCTCATGGAAGCCCTCAAGAG-----ACGGAT 71999

QY 122 GlnHisLeuMetCysGlnGluHisGluAspGluLysIleAsnIleTyrCysLeuSerCys 141

Db 72000 CAAAGAAATGTCATGTGAGGACAC---GGAGACACTTCACCTGTCTTCGAAAGCAG 72056

QY 142 GluValProThrCysSerLeuCysLysValPheGlyAlaHisLys 156

Db 72057 GGGCAGCTCATCTGCTGGCGCTGTGAGCGGACACACAGACAA 72101

RESULT 12

US-08-724-394A-21

Sequence 21, Application US/08724394A

GENERAL INFORMATION:

APPLICANT: Feder, John N.

APPLICANT: Krommal, Gregory S.

APPLICANT: Lauert, Peter M.

APPLICANT: Ruddy, David A.

APPLICANT: Thomas, Winston

APPLICANT: Tsuchihashi, Zenta

APPLICANT: Wolff, Roger K.

TITLE OF INVENTION: Megabase Transcript Map: No. 5872237e1

TITLE OF INVENTION: Sequences and Antibodies Thereof

NUMBER OF SEQUENCES: 31

CORRESPONDENCE ADDRESSES:

ADDRESS: TOWNSEND and TOWNSEND and CREW LLP

STREET: Two Embarcadero Center, 8th Floor

CITY: San Francisco

STATE: CA

COUNTRY: USA

ZIP: 94111-3834

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/724,394A

FILING DATE: 01-OCT-1996

CLASSIFICATION: 536

ATTORNEY/AGENT INFORMATION:

NAME: Fitts, Renee A.

REGISTRATION NUMBER: 35,136

REFERENCE/DOCKET NUMBER: 017957-000100

## TELECOMMUNICATION INFORMATION:

TELEPHONE:	415-576-0200
TELEFAX:	415-576-0300
INFORMATION FOR SEQ ID NO:	21
SEQUENCE CHARACTERISTICS:	
LENGTH:	246240 base pairs
TYPE:	nucleic acid
STRANDEDNESS:	not relevant
TOPOLOGY:	not relevant
MOLECULE TYPE:	CDNA

FEATURE:

NAME/KEY: misc.feature

LOCATION: 1..246240

OTHER INFORMATION: /note= "HLA-H. CONTIG"

US-08-724-394A-21

Alignment Scores:

Pred. No.:	0.0753	Length:	246240
Score:	131.00	Matches:	37
Percent Similarity:	42.96%	Conservative:	21
Best Local Similarity:	27.41%	Mismatches:	53
Query Match:	6.85%	Indels:	24
DB:	2	Gaps:	6

US-09-908-988b-2 (1-366) x US-08-724-394A-21 (1-246240)

QY 22 LysGlnLeuIleCysProIleCysLeuGluMetPheSerLysProValIleLeuPro 41

Db 71769 GAGGAAGCACCCTGCTCCATTCGCTGAGCCCTGATGACCAACCAAGCAATC---AAC 71825

QY 42 CysGlnHisAsnLeuCysArgLysCysAlaAsnAspValPheGlnIleAsnProLeu 61

Db 71826 TGTGACACAGCTACTGCTGCTGATATACAGACTTCTTAA-----AACCCAGC 71879

QY 62 TrpGlnSerArgLysThrValSerSerGlyArgPheArgCysProSerCys 81

Db 71880 CAAAGCAACGACGACGACGAC-----TTCTGCTGTCTCCAGCTGT 71921

QY 82 ArgHisGluValIleuAspArgHisGlyValTyrGlyLeuGlnArgAsnLeuVal 101

Db 71922 CGGCTCCATTCATATATGAT-----AGCCTCGACCAACAAGCAGCTG 71966

QY 102 GluAsnIleIleAspIleTyrLysGlnGluSerSerArgProLeuHisAlaLysAlaGlu 121

Db 71967 GGAAGCCTCATGGAAGCCCTCAAGAG-----ACGGAT 71999

QY 122 GlnHisLeuMetCysGlnGluHisGluAspGluLysIleAsnIleTyrCysLeuSerCys 141

Db 72000 CAAAGAAATGTCATGTGAGGACAC---GGAGACACTTCACCTGTCTTCGAAAGCAG 72056

QY 142 GluValProThrCysSerLeuCysLysValPheGlyAlaHisLys 156

Db 72057 GGGCAGCTCATCTGCTGGCGCTGTGAGCGGACACACAGACAA 72101

RESULT 13

US-08-724-394A-22

Sequence 22, Application US/08724394A

Patent No. 5872237

GENERAL INFORMATION:

APPLICANT: Feder, John N.

APPLICANT: Krommal, Gregory S.

APPLICANT: Lauert, Peter M.

APPLICANT: Ruddy, David A.

APPLICANT: Thomas, Winston

APPLICANT: Tsuchihashi, Zenta

APPLICANT: Wolff, Roger K.

TITLE OF INVENTION: Megabase Transcript Map: No. 5872237e1

TITLE OF INVENTION: Sequences and Antibodies Thereof

NUMBER OF SEQUENCES: 31

CORRESPONDENCE ADDRESSES:

ADDRESS: TOWNSEND and TOWNSEND and CREW LLP

STREET: Two Embarcadero Center, 8th Floor

CITY: San Francisco



```

OY 123 sLeuMetCysGluGlnHisGluAspGluLysIleAsnIleTyrCysLeuSerCysGluVal 143
Db 694 -----ATGAGAGCTCTTCTGCGACAGACCAGCAGAC 722
OY 143 lProThrCysSerLeuCysLysValPheGlyAlaHisLysAspCysGluValAlaProle 163
Db 723 CTGCATCTGCTACCTTTGC---ATGTTCCAGAGCACAAGATCATAGACCGCTGACAGT 779
OY 163 uProThrIleTyrLysArgGlnLysSerGlnLeuSerAspGlyIleAlaMetLeuValAl 183
Db 780 GAGAGAGCCAGCCAGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 818
OY 183 aGlyAsnAspArgValGlnAlaValIleThrGlnMetGluGlnValCysGln-----CT 200
Db 819 GCAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 878
OY 201 -----ThrIleGluAspAsnSerArgArgGlnLysGlnLeuLeuAsnGlnArg 216
Db 879 GAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 938
OY 216 gPheGluThrLeuCysAlaValIleGluGluGluArgLysGluGluLeuGlnAlaLeuAl 236
Db 939 CTTCGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 998
OY 236 aArgGluGlnGluGluLysLeuGlnArgValArgGlyLeuLeuLeuArgGlnTyrGlyAspH 256
Db 999 GAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1046
OY 256 sLeuGluGlySerSerLysLeuValGluSerAlaIleGlnSerMetGluGlu----- 273
Db 1047 TCTGGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1106
OY 274 ----ProGlnMetAlaLeuTyrLeuGlnGlnAlaLysGlnLeuLeuAsnLysValGlyAl 292
Db 1107 CATCAGCAGACTGTGTGTTCTGTCGAGAA-----TTTGGTGC 1145
OY 292 aMetSerLysValGlnLeuAlaGlyArgProGluProGlyTyr----- 306
Db 1146 ATTTGATGAGCAATTACTCTCTCCGCCGCCATCATCATGCTCTGTCGAGAGG 1205
OY 307 -----GluSerMetGluGlnPhe-----SerValSerValG 317
Db 1206 GAGAGGCGCTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1265
OY 317 uHisValAlaGluMetLeuArg 324
Db 1266 CCACGTTGAGAGAGATGTGCAAG 1287

```

RESULT 15

PCT-US93-05794-2  
Sequence 2, Application PC/TUS9305794  
GENERAL INFORMATION:

APPLICANT: The Regents of the University of California  
TITLE OF INVENTION: Gene for Ataxia-telangiectasia  
NUMBER OF SEQUENCES: 45  
CORRESPONDENCE ADDRESS:

ADDRESSEE: Leona L. Lauder  
STREET: 177 Post Street, Suite 800  
CITY: San Francisco  
STATE: California  
COUNTRY: San Francisco  
ZIP: 94108-4731

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US93/05794  
FILING DATE: 19930618

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/903,466

```

FILING DATE: 22-JUN-1992
ATTORNEY/AGENT INFORMATION:
NAME: Lauder, Leona L.
REGISTRATION NUMBER: 30,863
REFERENCE/DOCKET NUMBER: 91-077-1 PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-421-4973
TELEFAX: 415-421-1663
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1767 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1764
PCT-US93-05794-2

Alignment Scores:
Pred. No.: 7,74e-05
Score: 127.00
Percent Similarity: 40.918
Best Local Similarity: 21.438
Query Match: 6.64%
DB: 5
Gaps: 13

US-09-908-988b-2 (1-366) x PCT-US93-05794-2 (1-1767)
OY 46 LeuCyArgLysCysAlaAsnAspValPheGlnAla-SerAsnProLeuTyrGlnSerArg 65
Db 514 CTGTGGAGACTCTGTGATCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 573
OY 65 g---GlySerThrThrValSerSerGlyArgPheArgCysProSerGysArg---H 83
Db 574 GCCTCCTTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 632
OY 83 sGluValAlaLeuAspArgHisGlyValTyrGlyLeuGlnArgAsnLeuValGluAs 103
Db 633 CCAG-----CTGCTCGAGCC 647
OY 103 nIleIleAspIleTyrLysGlnLysSerSerArgProLeuHisAlaLysAlaGluGlnH 123
Db 648 CATCGGAGACTTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 693
OY 123 sLeuMetCysGluGlnHisGluAspGluLysIleAsnIleTyrCysLeuSerCysGluVal 143
Db 694 -----ATGAGAGCTCTTCTGCGACAGACCAGCAGAC 722
OY 143 lProThrCysSerLeuCysLysValPheGlyAlaHisLysAspCysGluValAlaProle 163
Db 723 CTGCATCTGCTACCTTTGC---ATGTTCCAGAGCACAAGATCATAGACCGCTGACAGT 779
OY 163 uProThrIleTyrLysArgGlnLysSerGlnLeuSerAspGlyIleAlaMetLeuValAl 183
Db 780 GAGAGAGCCAGCCAGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 818
OY 183 aGlyAsnAspArgValGlnAlaValIleThrGlnMetGluGlnValCysGln-----CT 200
Db 819 GCAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 878
OY 201 -----ThrIleGluAspAsnSerArgArgGlnLysGlnLeuLeuAsnGlnArg 216
Db 879 GAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 938
OY 216 gPheGluThrLeuCysAlaValIleGluGluGluArgLysGluGluLeuGlnAlaLeuAl 236
Db 939 CTTCGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 998
OY 236 aArgGluGlnGluGluLysLeuGlnArgValArgGlyLeuLeuLeuArgGlnTyrGlyAspH 256

```

[illegible]

Percent Similarity:	40.91%	Conservative:	60
Best Local Similarity:	21.43%	Mismatches:	104
Query Match:	6.64%	Indels:	79
DB:	1	Gaps:	13
US-09-908-988B-2 (1-366) x US-07-903-466-1 (1-3018)			
QY 46	LeuCYsArGlyScyAlaLAsnAPrAlaPheGlnAla-SerAsnProLeuThrGlnSerAr	65	
DB 638	CTGTGCGAGCTCTGATCGAGCGCAACAGACAGAGGGGGTGAAGTCTGCTGGTGTGCCAG	697	
QY 65	G--GlySerThrPheValSerSerGlyAlaGrPheArgCysProSerCysArg--H1	83	
DB 698	GGCTGCTTGTGGAGACTGATGATCTCAAGCCCAAGCCCACTGGAGGGGGCGGCTTC--GAGAGCA	756	
QY 83	sGLuValAlaLeuAspArgHISGLyAlaTyrGLyLeuGlnArgAsnLeuValGLuAs	103	
DB 757	CCAG-----CTGCTCGAGCC 771		
QY 103	AlleleAspLleTyrLysGlnGlnSerSerArGrProLeuHISAlaLysAlaGLuGlnH1	123	
DB 772	CATCGGGAGCTTTGAGGCCCGCAAGGT-----CCGGGCAATGCGCAAGCA-----	817	
QY 123	sLeuMetCysGlnGlnHISGlnAspGlnLysLeuSLeaSNleTyrCysLeuSerCysGLuVal	143	
DB 818	-----ATGAGAGCTTCTTGGCAGACGACGACAGAC 846		
QY 143	lProThrCysSerLeuCysLysValaPheGLyAlaHISLysAspCysGLuValaLProLe	163	
DB 847	CTGCATCTGCTACCTTGGC--ATGTTCCAGAGGACACAAATCATATGACCCCTGACAGT	903	
QY 163	uProThrLleTyrLysArgGlnLysSerGlnLeuSerAspGlyLleAlaMetLeuValaL	183	
DB 904	GAGAGAGGCCAACGCCGAGAGAGAGAGAGAGAGCTGTA-----CT 942		
QY 183	agLysAsnAspArgValaGlnAlaValLleThrGlnMetGLuGLuValaCysGln-----	200	
DB 943	GCAAAGGAGACACTGCTGACGTCAAGTCAATGATGATGAGATGAGAGCTGAGAAAGTGCA	1002	
QY 201	-----ThrlleGlnAspAsnSerArGrArgGlnLysGlnLeuLeuAsnGlnAr	216	
DB 1003	GAGAGAGAGAGACCGGATCATCAAGAGCTTACACCAACATGAGAAAGGCCATCTGGAGCAGAA	1062	
QY 216	grPheGlnThrLeuCysAlaValaLeuGlnGlnArgLysGLyLleuLeuGlnAlaLeuAl	236	
DB 1063	CTTCCGGGAGACTGTGGCGGAGCTGGAGAAAGCAAAAGAGAAAGTAGGGCTGGCTGGA	1122	
QY 236	aArgGlnGlnGlnLysLeuGlnArgValaArgGlyLeuLleArgGlnTyrGLyAspH1	256	
DB 1123	GCAGCGGAGACAGAGATGCTGTGGACAAATGAGAGGTATGATG-----GATGC 1170		
QY 256	sLeuGlnGlySerSerLysLeuValaGlnSerAlaLleGlnSerMetGLuGlu-----	273	
DB 1171	TCTGGATGAGAGAGCCCAAGGTGCTGCTAGAGACAAACAGACCCGGAGAGAGCTGATAG	1230	
QY 274	-----ProGlnMetAlaLeuArgLysGlnGlnAlaLysGLuLeuLleAsnLysValaGLyAl	292	
DB 1231	CATCAGGCACTGTGTGTGTTGTTCTGACAGAA-----TTTGGTGC 1269		
QY 292	ameLysLysValaGlnLeuAlaGLyArgProGlnProGlyLyr-----	306	
DB 1270	ATTGATATGAGCAATTACTCTCTCCCGCCACCTGCCACATATGATCTCTGTGAGAGG 1329		
QY 307	-----GluSerMetGlnGlnPhe-----SerValSerValGL 317		
DB 1330	GGAAGGCGCTGGAGACAGTACATAGGCAATTCAGAGAGACAGACCTGCTCAATGTATGATCATCG	1389	
QY 317	uHISValaLleGlnMetLeuArg 324		
DB 1390	CCAGCTTGAGAGATGTGCAG 1411		

```

: Sequence 1, Application PC/TUS9305794
: GENERAL INFORMATION:
: APPLICANT: The Regents of the University of California
: TITLE OF INVENTION: Gene for Ataxia-Telangectasia
: TITLE OF INVENTION: Complementation Group D (ATDC)
: NUMBER OF SEQUENCES: 45
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Leona L. Lauder
: STREET: 177 Post Street, Suite 800
: CITY: San Francisco
: STATE: California
: COUNTRY: San Francisco
: ZIP: 94108-4731
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: PCT/US93/05794
: FILING DATE: 19930618
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/903,466
: FILING DATE: 22-JUN-1992
: ATTORNEY/AGENT INFORMATION:
: NAME: Lauder, Leona L.
: REGISTRATION NUMBER: 30,863
: REFERENCE/DOCKET NUMBER: PCT
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 415-421-4973
: TELEFAX: 415-421-1663
: INFORMATION FOR SEQ ID NO. 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 3018 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA
: HYPOTHEICAL: NO
: ANTI-SENSE: NO
: PCT-US93-05794-1

Alignment Scores:
Pred. No.: 0.000183 Length: 3018
Score: 127.00 Matches: 60
Percent Similarity: 40.91% Conservative: 66
Best Local Similarity: 21.43% Mismatches: 104
Query Match: 6.64% Indels: 79
DB: 5 Gaps: 13

US-09-908-988B-2 (1-366) x PCT-US93-05794-1 (1-3018)
QY 46 LeucysArglyscyAlasnaAspyalPheGlnAla-SerAsnProLeuTyrpgInserAr 65
||| ||| :: :||| ||| ||| ||| :|||
Db 638 CTGTGCAGACTCCTGGATGGCAACAGAGAAGCGGTCAGTGCTCTGCTGTGCCAG 697
65 g--gLySerThrTrValSerSergLyglYarGpheaRgcysProSerCysArg---Hl 83
||| ||| :: :||| ||| :||| ||| ||| |||
Db 698 GCCCTCTCTGTGGACTCATCTCAAGCCACCCTGGAGGCGCGGCTTC-CGAGACA 756
83 sgluValleuAsparghISglYalTyrclYleuGlnArgasnLeuValGuas 103
:::
Db 757 CCAG-----CTGCTCGAGCC 771
QY 103 llelleaspIleTyrlYsglngluSerSerArgProleuhIsAlalysAlagluGlnHl 123
||| ||| :||| :||| ||| |||
Db 772 CATCGGGAGACTTGAAGGCCCAAGTGT-----CCGTGCATGGCAAGAGC----- 817
123 sleuMcCysglugluIHISgluaSpGlulYsIlEasnIleTyrcYsleuSerCysgluVa 143
818 -----ATGAGAGCTCTTCTGCCAGACCGAACCGACAGAC 846
143 lprothrCysSerleuCySlYsValPheGluAlahnlSYsasPCysgluValAlaIPole 163
```

```

Db      847  CTGCATCTCTCACTTTCG---ATGTCACGAGGACGACAAATCATAGCACCGTCACT 903
              ||| ||||| :||| |||||:|
QY      163  uprothrlleTyrlYrsArlGlnLysSerGlnLeuSerAspGlylleAlaMetLeuValAl 183
              : :|||:|
Db      904  GGAGGAGGCCAAGGCCGAGAGAGAGAGAGAGCTGTCA-----CT 942
              :|||:|
QY      183  aGlyAsnaSPaTgValGlnAlaVallleThrlGlnMetGlnValAlCysGln----- 200
              :|||:|
Db      943  GCMAAAGGAGCGCTGCTCACTCAAGATCTTAGAGATTGGAGTGAAGCTGAAGATGGCA 1002
              :|||:|
QY      201  -----ThrlleGlnuSPaSPaSerArgrGlnLysGlnLeuLeuAsnGlnAr 216
              :|||:|
Db      1003  GAAGGAGAGAGACCCGATCAACGCTTCACACCAATGAGAAAGGCCATCTGTGAGACAGA 1062
              :|||:|
QY      216  gPheGlnThrlLeuCySAlaValLeuGlnLysGlnLysGlnLeuLeuGlnAlaLeuAl 236
              :|||:|
Db      1063  CTTCGCGAGCTGTGCTCCGGGACCTGGAGAAAGCAAAAGAGAGAGAGGCTGCGCTGGA 1122
              :|||:|
QY      236  aArGlnGlnGlnGlnLysLysLeuGlnArgrValArgrGlyLeuLeuArgrGlnTyrGlyAspH 256
              :|||:|
Db      1123  GGAGGGGGAGCAGAGATGCTGTGACCAATGAAATGAGTGAATGATC-----GATCC 1170
              :|||:|
QY      256  sLeuGlnGlySerSerLysLeuValGlnSerAlalleGlnSerMetGlnGlu----- 273
              :|||:|
Db      1171  TCTGTGATGAGAGAGCCAGAGCTGCTCATGAGAGACAAGCACCGGAGACAGCTCATAG 1230
              :|||:|
QY      274  ----ProGlnMetAlaLeuTyrLeuGlnGlnAlaLysGlnLeuIleAsnLysValGlyAl 292
              :|||:|
Db      1231  CATCAACGCACTGTGTTGTCTCGAGAA-----TTTGTGTC 1269
              :|||:|
QY      292  aMetSerLysValGlnLeuAlaGlyArgrProGlnProGlyTyr----- 306
              :|||:|
Db      1270  ATTGATGAGCAATTACTCTCTCCCCCAACCCCTGCCACACATATCATGTGCTGTGAGGG 1329
              :|||:|
QY      307  -----GlnSerMetGlnGlnPhe-----SerValSerValGln 317
              :|||:|
Db      1330  GGAGGGCCCTGGCAGGTCACTAGGCAACCTCAAGAGCAGACCTGCTCAATGATATGATCGC 1389
              :|||:|
QY      317  uHisValAlaGlnMetLeuArg 324
              :|||:|
Db      1390  CCACGTTGAGAAAGATGTGCAAG 1411
              :|||:|

RESULT 18
US-08-801-308-2
; Sequence 2, Application US/08801308
; Patent No. 6368790
; GENERAL INFORMATION:
; APPLICANT: Scott, Robert E.
; TITLE OF INVENTION: CDNA ENCODING P2P PROTEINS AND USE OF
; TITLE OF INVENTION: P2P CDNA-DERIVED ANTIBODIES AND ANTISENSE REAGENTS IN
; TITLE OF INVENTION: DETERMINING THE PROLIFERATIVE POTENTIAL OF NORMAL,
; TITLE OF INVENTION: ABNORMAL AND CANCER CELLS IN ANIMALS AND HUMANS
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Weiser & Associates, P.C.
; STREET: 230 S. Fifteenth Street, Suite 500
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19102
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentln Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/801,308
; FILING DATE: 18-FEB-1997
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Weiser, Gerard J.
; REGISTRATION NUMBER: 19,763

```







Query Match: 6.41% Indels: 109  
DB: 4 Gaps: 19  
US-09-908-988b-2 (1-366) x US-08-697-610-1 (1-2240)

QY 20 LeuGlulysGlnLeuIleCysProIleCysLeuGlnMetPheSerLysProValIle 39  
DB 275 GTGGAGGACAGATACAGATGAGAGAGTCCAGCTGCTGTGCAGCCGAGAGCAG-- 331  
QY 40 LeuProCysGlnHisAsnLeuCysArgLysCysAlaAsnAspValPheGlnIleAsn 59  
DB 332 ACCGAGTGTGGGACCCCTCTCTCGAGAGCTGCATGGCGCCCTCTGAGCTTCAACT 391  
QY 60 ProLeuTrpGlnSerArgLysSerThrThrValSerSerGlyArgPheArgCysPro 79  
DB 392 CCA-----AAATGTACA 403  
QY 80 SerCysArgHisGluValIleLeuAspArgHisGlyValTyrGly-----LeuGln 96  
DB 404 GCGTGTCAAGAGAGCATCGTTAAAGATAG-----GTTTAAAGATTAATGCTGCAAG 457  
QY 97 ArgAsnLeuLeuValGluAsnIleIleAspIleTyrLysGlnGlnSerSerArgProLeu 116  
DB 458 AGAGAATTTCTGCT-----CTTCAGATCTATTGTCCGAATGAAGACAGA----- 502  
QY 117 HisAlaLysAlaGluGln-----HisLeuMetCysGluGlu 128  
DB 503 ---GGTTGTGCAGAGCATTAACGCTGGACATCTGCTGCTCAATTTAAATAATGATTGC 559  
QY 129 HisGluAspGluLysIleAsnIleTyrCysLeuSerCysGluValPro----- 144  
DB 560 CATTTTGAAGAA-----CTTCCATGTGTGCGCTCGCTGCAAGAA 601  
QY 144 ----- 144  
DB 602 AAGGCTTGTAGAAAGACTGCGAGACCATGTGAGAGAGGCTGTAATACCGGAGAACCC 661  
QY 145 ThrCysSerLeuCysLys-----ValPheGlyAlaHisLysAsp--- 157  
DB 662 ACATGAGCAGCAGCAAGATCAGGTTCCGATGATCGCGTGCAGAAACAGAGACACCC 721  
QY 158 -----CysGluValAlaProLeuProThrIleTyrLysArgGlnLysSerGluLeu 174  
DB 722 GACTGTCCTGCTGCTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 781  
QY 175 SerAspGlyIleAlaMetLeuValAlaGlyAsnAspArgValGlnAlaVal----- 191  
DB 782 ACCGAGGGGACAAACAGCATCAAGAGGCCAGAGGCCAGCTCCGCCCTGCAGACATC 841  
QY 192 -----IleThrGlnMetGluGluValCysGlnThrIleGluAspAsn 205  
DB 842 AACCTGCTGAAGAGTGGAGCAACTGCTGCAAAAGAGGTTCTTGTGCAAGATGAA 901  
QY 206 SerArgArgGlnLysGlnLeuLeuAsnGlnArgPheGlnThrLeuGlyAla-----Val 223  
DB 902 AGTGTAGAAAAAACAAGACATACAAAGTTTGCACATCAGATATGTACCTTTGAATTT 961  
QY 224 LeuGluGluArgLysGlyLysLeuLeuGlnAlaLeuAlaArgGlnGlnGlnGlnLysLeu 243  
DB 962 GAAATTTGAGAGCAAAAGAAATGCTT-----CGAATATATGATCCAAATC 1009  
QY 244 GlnArgValArgGlyLeuIleArgGlnTyrGlyAspHisLeuGlnGlySerSerLysLeu 263  
DB 1010 CTTCAATTTACACGAGTGTATACACACCAAGCAAGAGAACTGAGAGGCTTGCACAAAGAG 1069  
QY 264 ValGluSerAlaIleGlnSerMetGlnGluProGlnMetAlaLeuTyrLeuGlnGlnAla 283  
DB 1070 ATCCGCTCTCCGAGAGACACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1123  
QY 284 LysGluLeuIleAsnLysValAlaMetSerLysValGlu-----LeuAlaGlyArgPro 302  
DB 1124 GAGTCCCTCCAGAACCGGCTGACCGAGTGTGAGAGAGCGTGTGACAAAGCGCGGAGCAAGTG 1183

QY 303 GluProGlyTyrGlnSerMetGlnGlnPheSerValSerValGluHisValAlaGluMet 322  
DB 1184 GCTCGAAGACAGAGCCTGTGAGATCC-----CAGCTGAGCCGAGCATGACAGATG 1234  
QY 323 LeuArgThrIleAspPheGlnProGlyAlaAlaGlyAspGluGluAspAspMetAla 342  
DB 1235 CTGAGTGTGACAGCATCCG-CCTAGCCGACATGACCT-----GCG 1275  
QY 343 LeuAspGlyGlu-----GlnGlyAsnAlaGlyLeuGluGluGluArgLeu 357  
DB 1276 CTTCCAGGCTCTCGAGAGACCCGACCTACAAATGAGTGTCTGGAAGATTCGGACTA 1335  
QY 358 AspValProGluGlnSerGlyLeuHis 366  
DB 1336 CAAGCGCGGAGAGAGAGCGCTCAT 1362

RESULT 20  
US-08-349-357-1  
; Sequence 1, Application US/08349357  
; Patent No. 6265556  
; GENERAL INFORMATION:  
; APPLICANT: Reed, John C.  
; APPLICANT: Sato, Takaaki  
; TITLE OF INVENTION: CD40 Associated Proteins  
; NUMBER OF SEQUENCES: 17  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Campbell and Flores  
; STREET: 4370 La Jolla Village Drive, Suite 700  
; CITY: San Diego  
; STATE: California  
; COUNTRY: USA  
; ZIP: 92122

COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/349,357  
; FILING DATE: 02-DEC-1994  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Campbell, Cathryn A.  
; REGISTRATION NUMBER: 31,815  
; REFERENCE/DOCKET NUMBER: P-LJ 1203  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (619) 535-9001  
; TELEFAX: (619) 535-8949  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2240 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 137..1766  
; US-08-349-357-1

Alignment Scores:  
Pred. No.: 0.000378 Length: 2240  
Score: 122.50 Matches: 86  
Percent Similarity: 44.7% Conservative: 55  
Best Local Similarity: 21.03% Mismatches: 160  
Query Match: 6.41% Indels: 109  
DB: 4 Gaps: 19

US-09-908-988b-2 (1-366) x US-08-349-357-1 (1-2240)

QY 20 LeuGluLysGlnLeuIleCysProIleCysLeuGlnMetPheSerLysProValIle 39  
DB 275 GTGGAGGACAGATACAGATGAGAGAGTCCAGCTGCTGTGCAGCCGAGAGCAG-- 331



GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: December 3, 2002, 12:12:18 ; Search time 261 Seconds  
(without alignments)  
3157.975 Million cell updates/sec

Title: US-09-908-988b-2  
Perfect score: 1912  
Sequence: 1 MNVTGFKPLGDHNMNDL.....EGNAGLEERLDEVEGSLH 366

Scoring table:  
BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2185239 segs, 112599159 residues  
Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=framed\_p2n\_model -DEV=xlh  
-O=/cgn2\_1/USPTO.spool/US09908988/runat\_26112002\_112153\_11206/app.query.fasta\_1.519  
-DB=N\_geneseq\_101002 -OVRT=fastap -SUFFIX=trng -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blsum62 -TRANS=human40.cdi  
-LIST=45 -DOCMALIG=200 -THR.SCORE=pct -THR.MAX=100 -THR.MIN=0 -ALIGN=20  
-MODE=LOCAL -OUTFILE=plo -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US09908988.ecgn.1.1.0.ernat.26112002\_112153\_11206 -NCPU=6 -ICPU=3  
-NO\_XLPHY -NO\_MMP -FARGOUDRY -NEG\_SCORES=0 -MATT=LONGLOG -DEV=TIMEOUT-120  
-WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6 -Fgapext=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :  
1: N\_Geneseq\_101002.\*  
2: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.\*  
3: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.\*  
4: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.\*  
5: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.\*  
6: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.\*  
7: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.\*  
8: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.\*  
9: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.\*  
10: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.\*  
11: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.\*  
12: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.\*  
13: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.\*  
14: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.\*  
15: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.\*  
16: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.\*  
17: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.\*  
18: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.\*  
19: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.\*  
20: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.\*  
21: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.\*  
22: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.\*  
23: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.\*  
24: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.\*  
25: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1912	100.0	1431	24	ABA99061
2	1780	93.1	1913	22	AAH68563
3	1780	93.1	1990	22	AAH78026
4	1725	90.2	1349	21	AAH72433
5	1599	83.6	2040	24	ABN85313
6	1261.5	66.0	1500	22	AAE27653
7	1162	60.8	2590	24	ABA99062
8	1147.5	58.1	1762	22	AAE25855
9	1110	58.1	1597	24	ABA99063
10	1101.5	57.6	1210	22	AAE42407
11	1098.5	57.5	1231	22	AAH80037
12	1098.5	57.5	1764	22	ABA83058
13	1091.5	57.1	1781	22	AAE25842
14	987.5	51.6	1183	22	AAH90117
15	920.5	48.1	1796	22	AAH89924
16	659.5	34.5	531	21	AACT5285
17	659.5	34.5	531	21	AACT5285
18	614	32.1	587	24	ABN78104
19	427.5	22.4	4345	23	AAE26314
20	427.5	22.4	4345	23	ABV22250
21	427.5	22.4	4345	23	ABV23763
22	427.5	22.4	4345	23	ABV24205
23	427.5	22.4	4345	23	ABV24320
24	427.5	22.4	4345	23	ABV24652
25	427.5	22.4	4345	23	ABV27751
26	427.5	22.4	4345	23	ABV28087
27	427.5	22.4	4345	23	ABV29637
28	416	21.8	446	22	AAE26303
29	411	21.5	2058	23	AAE82247
30	411	21.5	2237	23	AAE82248
31	411	21.5	6632	22	AAE82234
32	306	16.0	650	21	AACT5760
33	298	15.6	391	21	AACT5969
34	286	15.0	2272	23	AAE77725
35	275	14.4	1676	22	AAE44635
36	274	14.3	1409	22	ABQ11959
37	274	14.3	1755	21	AAE16294
38	274	14.3	1801	22	AAE18517
39	274	14.3	2565	23	ABV25262
40	259.5	13.6	1761	23	AAE85165
41	259	13.5	1276	24	ABO72500
42	255.5	13.4	3262	21	AAE83838
43	234	12.2	2981	21	AACT7707
44	225	11.8	599	22	AAE16184
45	225	11.8	599	22	AAE16185

## ALIGNMENTS

RESULT 1	ABA99061	standard; DNA; 1431 BP.
ID	ABA99061	
AC	ABA99061	
XX		
DT	15-JUL-2002	(first entry)
XX		
DE	Murine muscle ring finger protein 1 (MURF-1) coding sequence.	
XX		
KW	Muscle ring finger; MURF-1; mouse; cardiant; microtubule;	
KW	intermediate filament; striated muscle; cardiac hypertrophy;	
KW	heart disease; gene; ds.	
XX		
OS	Mus musculus.	
XX		
Key	Location/Qualifiers	
FT	CDS	199..1299

```

FT      /tag="a
XX      /product="MURF-1"
XX      MO200206318-A2.
XX      PD 24-JAN-2002.
XX      PF 18-JUL-2001; 2001WO-US22896.
XX      PR 18-JUL-2000; 2000US-219020P.
XX      PA (TEXA ) UNIV TEXAS SYSTEM.
XX      PI Olson EN, Spencer JA;
XX      DR WPI: 2002-241506/29.
XX      PT P-PSDB; ABB08275.
XX      PS Novel muscle ring finger protein useful for drug screening, and for
        diagnosing and treating diseases, particularly cardiomyopathies -
        Claim 4; Page 123-125; 134pp; English.

```

```

XX      The sequence encodes murine muscle ring finger protein 1 (MURF-1). The
XX      invention relates to a purified muscle ring finger (MURF) protein,
XX      selected from MURF-1, MURF-2 and MURF-3. The MURF proteins of the
XX      invention are involved in microtubule and intermediate filament
XX      stabilisation of striated muscle cells and have cardiant activity.
XX      The MURF proteins are useful for screening a candidate substance
XX      for MURF protein-binding activity, in a cell, cell-free system or
XX      in vivo, and its effect on interaction of MURF with microtubules,
XX      homodimerisation of MURF, MURF-1, MURF-2 or MURF-3 stabilisation
XX      of microtubules, interaction of MURF with intermediate filaments,
XX      e.g. desmin, vimentin and cytokeratin, and heterodimerisation of MURF.
XX      The screened compounds are useful for treating and preventing of
XX      cardiac hypertrophy and heart diseases. MURF proteins are useful as
XX      antigens to immunise animals for the production of antibodies.

```

Sequence 1431 BP; 338 A; 384 C; 473 G; 236 T; 0 other;

```

XX      Alignment Scores:
XX      Pred. No:      1.86e-178      Length:      1431
XX      Score:          1912.00      Matches:      366
XX      Percent Similarity: 100.00%      Conservative: 0
XX      Best Local Similarity: 100.00%      Mismatches: 0
XX      Query Match:      100.00%      Indels:      0
XX      DB:                24      Gaps:          0

```

US-09-908-988b-2 (1-366) x ABA99061 (1-1431)

```

XX      1 MetAspPheThrValGlyPheLysProLeuLeuGlyAspAlaHisAsnMetAspAsnLeu 20
XX      199 ATGAAGCTTCACGCGTGGGTTTCAAGCGCGCTGCTAGGGGATCGGCACACATGTGAACACTTG 258
XX      21 GluLysGlnLeuIleCysProIleCysLeuGluMetPheSerLysProValIleLeu 40
XX      259 GAGAGCAGCATCTATTTGCCCATCTGCTGAGAGATGTTTCCAAAGCCCGTGATCTTG 318
XX      41 ProCysGlnHisAsnLeuLysArgLysCysAlaAsnAspValPheGlnAlaSerAsnPro 60
XX      319 CCGTGTCCACACACAACTGTGGCGCAGGTGGCCACAGCATCTTCCAGGCTCTTAATCTCT 378
XX      61 LeuTPrGlnSerArgGlySerThrThrValSerSerGlyGlnArgPheArgCysProSer 80
XX      379 CTGTGGCAATCCCGGGGCTCCACACAGGCTGTCTTCAAGAGAGACGTTTCCGATGCCCATCT 438
XX      81 CysArgHisGlnValValLeuAspArgHisGlyValTyrGlyLeuGlnArgAsnLeu 100
XX      439 TGTAGCAGCAGAGGTGTCTCGACAGAGCATGTGTCTATGTGGCTCGACGGGAACCTGCTA 498
XX      101 ValGluAsnIleIleAspIleTyrLysGlnGluSerSerArgProLeuHisAlaLysAla 120
XX      499 GTGGAGAACATCATTTGACATCTACACAGCAGGAGTCTCCCGGGCCACTGCACGCCAAGGCT 558

```

```

XX      121 GluGlnHisIleuMetCysGlnGluHisGlnLysAspGluLysIleAsnIleTyrCysIleuSer 140
XX      559 GAAACAGCACTCATGTGTGAGACACGAGACGAGAAATCAACATCTACTGCTGAGC 618
XX      141 CysGlnValProThrCysSerLeuCysLysValPheGlnAlaHisLysAspCysGlnVal 160
XX      619 TGGAGAGTGGCCACCTGCTCTCTGCAAGGTTTTCGGCGCCGCCACAGAGACGTGAGGTG 678
XX      161 AlaProLeuProThrIleTyrLysArgGlnLysSerGluLeuSerAspGlyIleAlaMet 180
XX      679 GCCCTCTGGCCACCATTTACAAAGCCAGAAAGAGTGAAGTGAAGATGGCATGGCAGATG 738
XX      181 LeuValAlaGlnLysAsnAspArgValGlnAlaValIleTyrGlnMetGlnGlnValCysGln 200
XX      739 CTGTGGCGGGCAATGACCCGTGTGTCAGGACATGATCACCAGATGAGAGAGGTGTGCCAG 798
XX      201 ThrIleGluAspAsnSerArgArgGlnLysGlnLeuLeuAsnGlnArgPheGlnThrLeu 220
XX      799 ACCATTGAGGACAAACAGCCGACAGACAGAACCACTGTTAAACCAAGAGGTTCGAGACCCTG 858
XX      221 CysAlaValLeuGlnGluArgLysGlyGluLeuLeuGlnAlaLeuAlaArgGlnGlnIle 240
XX      859 TGGCGGCTTTTGGAGGAGCGCAGCGCAACGTGCTTCAAGCACTGGCCGGAGAGCAGAG 918
XX      241 GluLysLeuGlnArgValArgGlyLeuIleArgGlnTyrGlyAspHisIleuGlnLysSer 260
XX      919 GAAAGATTGCAAGCCGCGGCGGCTCTATCCGCGCATGACGGAGACCACTTGAAGGGCTCC 978
XX      261 SerLysLeuValGlnSerAlaIleGlnSerMetGlnGluProGlnMetAlaLeuTyrLeu 280
XX      979 TCAAGCTGTGTGAGTCCGCATCCAGTCCATGAGAGAGCCGCAATGGCTCTTCAACTTC 1038
XX      281 GlnGlnAlaLysGlnLeuIleAsnLysValGlnAlaMetSerLysValGlnLeuAlaGly 300
XX      1039 CACAGGCAAAAGGAGCGCATCAACAGAGTCCGGCGCAATGTCCAAAGTGAAGTGGCAGGA 1098
XX      301 ArgProGluProGlyTyrGlnLysSerMetGlnGlnPheSerValSerValGlnHisValAla 320
XX      1099 CGGCCGAGCCAGGCTATGAGAGCATGAGCAATTTCTGTGAGCGTGGAGACACGTGGGCC 1158
XX      321 GluMetLeuArgThrIleAspPheGlnProGlyAlaAlaGlyAspGlnLysAspAsp 340
XX      1159 GAAATGTTGGAACCAATCGACTTCCAGCGCGGCGCTGGGGATGAAGAGATGACGAC 1218
XX      341 MetAlaLeuAspGlyGlnGluGlyAsnAlaGlyLeuGlnGluLysArgLeuAspValPro 360
XX      1219 ATGGCTTTGGATGGGGAGAGGGAATGCGGGCTGGAGAGAGACCGGCTGGACGTGCCA 1278
XX      361 GluGlySerGlyLeuHis 366
XX      1279 GAAAGCTCAGGCGCTGCAC 1296
XX      DB 1279 GAAAGCTCAGGCGCTGCAC 1296
XX      RESULT 2
XX      AAH68563 standard; cDNA; 1913 BP.
XX      ID AAH68563;
XX      AC AAH68563;
XX      DT 13-SEP-2001 (first entry)
XX      DE Human protein HP03115 coding sequence.
XX      KW Human; gene therapy; tumour; ss.
XX      OS Homo sapiens.
XX      PN WO200142302-A1.
XX      PD 14-JUN-2001.
XX      PF 06-DEC-2000; 2000WO-JF08631.
XX

```

PR 06-DEC-1999; 99JP-0346863.  
 PR 06-DEC-1999; 99JP-0346864.  
 PR 08-FEB-2000; 2000JP-0031062.  
 PR 10-FEB-2000; 2000JP-0034090.  
 PR 10-FEB-2000; 2000JP-0034091.  
 PR 14-FEB-2000; 2000JP-0035829.  
 PR 14-FEB-2000; 2000JP-0035899.  
 PR 14-MAR-2000; 2000JP-0071161.  
 PR 30-MAY-2000; 2000JP-0160851.  
 PA (NISC-) JAPAN SCI & TECHNOLOGY CORP.  
 FI Kato S, Eguchi C, Saeki M.  
 XX WPI; 2001-381646/40.  
 DR P-PSDB; AAG93278.  
 XX  
 PT Human protein originated from tumor cell line, applicable as drug,  
 PT reagent for studying intracellular protein networks and protein source  
 PT for drug screening, also encoded cDNA for gene diagnosis and gene  
 PT therapy  
 PS Claim 3; Pages 236-240; 471pp; Japanese.  
 XX  
 CC The present sequence is a human protein coding sequence. The human  
 CC protein, preferably originated from tumor cell line, is applicable as a  
 CC drug, a reagent for studying intracellular protein networks and a protein  
 CC source for screening proteins for binding low molecular weight drugs. The  
 CC human protein coding sequence is useful for gene diagnosis and gene  
 CC therapy, expression vectors and transformant cells for detection of  
 CC ligands and receptors.  
 XX  
 SQ Sequence 1913 BP; 446 A; 537 C; 612 G; 318 T; 0 other;  
 Alignment Scores:  
 Pred. No.: 2.72e-165 Length: 1913  
 Score: 1780.00 Matches: 343  
 Percent Similarity: 97.51% Conservative: 10  
 Best Local Similarity: 94.75% Mismatches: 7  
 Query Match: 93.10% Indels: 2  
 DB: Gaps: 1  
 US-09-908-988b-2 (1-366) x AAH68563 (1-1913)  
 QY 1 MetasphenrValGlyPheLysProLeuLeuGlyAspAlaHisAsnMetAspAsnLeu 20  
 DB 303 ATGAACCTCAGACGTGGTTTCAAGCCGCTGCTAGGGAGATGCACACAGCATGGCAACCTG 362  
 QY 21 GluLysGlnLeuIleCysProIleCysLeuGluMetPheSerLysProValIleLeu 40  
 DB 363 GAGAAGCAGCTCATCTGCCCATCTGCTGGAGATGTCTCCAAACAGTGGATCCTG 422  
 QY 41 ProCysGlnHisAsnLeuCysArgLysCysAlaAsnAspValPheGlnAlaSerAsnPro 60  
 DB 423 CCCTGCCCAACACACTGCGCGCAAAATGTGCCAACGAGCTGCTCCAGGCGCTGAATCCT 482  
 QY 61 LeuTrpGlnSerArgLysSerThrValSerSerGlyLysArgPheArgCysProSer 80  
 DB 483 CTATGGCAGTCCCGGGGCTCCACCACTGTCTTCAAGAGCCGTTTCCGCTCCCATCG 542  
 QY 81 CysArgHisGlnValValLeuAspArgHisGlyValTyrGlyLeuGlnAlaArgAsnLeu 100  
 DB 543 TGCAGGCATGAGGTGTCTCTGACACAGACAGCGGTGTCTACGGCCTGCAAGAACTGTCTA 602  
 QY 101 ValGluAsnIleIleAspIleTyrLysGlnGluSerSerArgProLeuHisAlaLysAla 120  
 DB 603 GTGGAGAACATTATGACATTACAAAGCAGAGATCTCCAGGCGCTGCACTCCAAAGCT 662  
 QY 121 GluGlnHisLeuMetCysGlnGluGlnHisGlnAspGluLysIleAsnIleTyrCysLeuSer 140  
 DB 663 GAGCAGCAGCTCATGTGGGAGAGCATGAAGAAGAAATCATATTTTACTGCTGAGC 722  
 QY 141 CysGluValProThrCysSerLeuCysLysValPheGlyAlaHisLysAspCysGluVal 160

DB 723 TGTGAGTGCCCACTGCTCTCTGCAAGGTCTTCGGTGGCCACAAAGACATGTAGGTG 782  
 QY 161 AlAProLeuProThrIleTyrLysArgGlnLysSerGlnLeuSerAspGlyIleAlaMet 180  
 DB 783 GCCCAGTCCACCATTTTACAAAGCCAGAGAGTGAAGCTCAGCATGCGTCCGAGT 842  
 QY 181 LeuValAlaGlyAsnAspArgValGlnAlaValIleThrGlnMetGlnLysGln 200  
 DB 843 CTGTGGCAGGCAATGACCGCTGCAAGCAGTGAATCACAACATGAGAGAGGTGTCCAG 902  
 QY 201 ThrIleGluAspAsnSerArgArgGlnLysGlnLeuLeuAsnGlnArpPheGluThrLeu 220  
 DB 903 ACTATCGAGGCAAAATAGCCGGAGGCAAGACAGTGTGTTAAACAGAGGTTTGAAGCTG 962  
 QY 221 CysAlaValLeuGluGlnLysGlyGlyLeuLeuGlnAlaLeuAlaArgGlnGln 240  
 DB 963 TGCAGAGTGTGGAGAGCGCAAGGTTGAGTGTGCAAGCGCTGCGCGGAGCAAGAG 1022  
 QY 241 GluLysLeuGlnArgValArgGlyLeuIleArgGlnTyrGlyAspHisLeuGlnLysSer 260  
 DB 1023 GAGAACTGCAGCGCGCTCCGCGCTCATCCGTAGTATGGGACACCTGGAGGCTCC 1082  
 QY 261 SerLysLeuValGluSerAlaIleGlnSerMetGluGluProGlnMetAlaLeuTyrLeu 280  
 DB 1083 TCTAAGCTGGTGGAGTCTGCCATCCAGTCCATGGAAGACCAAAATGGCGCTGATCTC 1142  
 QY 281 GlnGlnAlaLysGlnLeuIleAsnLysValGlyAlaMetSerLysValGlnLeuAlaGly 300  
 DB 1143 CAGAGGCGCAAGGAGCTCATTAAGTCCGGGCGCATGTCCAGAGTGGAGTGGCAGG 1202  
 QY 301 ArgProGluProGlyTyrGluSerMetGluGlnPheSerValSerValGlnHisValAla 320  
 DB 1203 CGCCCGGAGCCAGCTTATGACAGCAATGAGCAATTCACCTAAGGTGGAGCACTGGCC 1262  
 QY 321 GluMetLeuArgThrIleAspPheGlnProGlyAlaAlaGlyAspGlnGluAspAsp 340  
 DB 1263 GAAATGCTGCGGACATGATGATCCAGCAGCGCTTCCGGGAGGAAAG-----GAG 1316  
 QY 341 MetAlaLeuAspGlyGlnGluGlyAsnAlaGlyLeuGlnGluGlnLysArgLysValPro 360  
 DB 1317 GTGGCCCGCAGGAGAGAGAGGCGCGCGCGGAGGAGAGAGCGGCGGATGGCCT 1376  
 QY 361 GluGly 362  
 DB 1377 TAAGCC 1382  
 RESULT 3  
 AAH78026  
 ID AAH78026 standard; cDNA; 1990 BP.  
 XX  
 AC AAH78026;  
 XX  
 DT 26-NOV-2001 (first entry)  
 XX  
 DE Nucleotide sequence of human cardiomyopathy associated protein (CAP).  
 XX  
 KW Human; cardiomyopathy associated protein; CAP; myocardial biopsy;  
 KW left ventricle assist device; LVAD; cardiomyopathy;  
 KW ventricular tachyarrhythmia; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key location/Qualifiers  
 FT CDS 283..1359  
 FT /\*tag= a  
 FT /product= "cardiomyopathy associated protein (CAP)"  
 XX  
 PD WO200162767-A1.  
 XX 30-AUG-2001.  
 PF 26-FEB-2001; 2001WO-US05888.

```

XX 24-FEB-2000; 2000US-0184825.
PR
XX (GENE-) GENE LOGIC INC.
XX
XX Bednarik D, Greene J, White M;
XX WPI; 2001-570621/64.
DR P-PSDB; AMG63832.
XX
XX Nucleic acid encoding a cardiomyopathy associated protein that is
XX differentially expressed in human left ventricle assist device (LVAD)
XX myocardial biopsy samples, useful for diagnosing, preventing or
XX treating cardiomyopathy -
XX
PS Claim 2; Page 49-51; 55pp; English.
XX
XX The present sequence encodes a human cardiomyopathy associated
XX protein (CAP) that is differentially expressed in human left ventricle
XX assist device (LVAD) myocardial biopsy samples. The predicted
XX isoelectric point of CAP is approximately 12.2. The CAP polynucleotide
XX and polypeptide are useful for diagnosing, preventing or treating
XX cardiomyopathy. The treatment diminishes the occurrence of at least
XX one of the following symptoms associated with cardiomyopathy: reduced
XX ejection fraction, increased left ventricular diastolic dimension,
XX decreased ventricular wall thickness, increased atrial size, valvular
XX regurgitation, exertional intolerance or ventricular tachyarrhythmia.
XX Modulators of CAP are also useful for preventing or treating
XX cardiomyopathy.
XX
XX Sequence 1990 BP; 524 A; 535 C; 613 G; 318 T; 0 other;
SQ
XX
XX Alignment Scores:
XX
XX   Alignment No.: 2, 87e-165      Length: 1990
XX   Score: 1780.00      Matches: 343
XX   Percent Similarity: 97.51%      Conservative: 10
XX   Best Local Similarity: 94.75%      Mismatches: 7
XX   Query Match: 93.10%      Indels: 2
XX                               Gaps: 1
XX
XX US-09-908-988b-2 (1-366) x AAH78026 (1-1990)
XX
XX 1 MetAsnPhetHrValGlyPheLysProLeuLeuGlyAspAlaHisAsnMetAspAsnLeu 20
XX 283 ATGAAGCTTCAGAGTGGGTTTCAAGCCGCTGAGGAGATGCACACAGCATGACAACTG 342
XX
XX 21 GlyLysGlnLeuLeuLysProLecLysLeuGlnMetPheSerLysProValValIleLeu 40
XX 343 GAGAAGCACCTCATCTGCCCCATCTGCTGAGATGTCTCCAAACAGTGTGATCTG 402
XX
XX 41 ProCysGlnHisAsnLeuLysCysArgLysCysAlaAsnAspValPheGlnAlaSerAsnPro 60
XX 403 CCTGCGCAACACACACTGTGTCGCCCAATGTGCGCAACAGCTTCTCCAGGCGCTCGAATCT 462
XX
XX 61 LeuTrpGlnSerArgLysSerThrThrValSerSerLysGlyArgPheArgCysProSer 80
XX 463 CTATGGACATCCCGGGGCTCCACACAGTGTCTTCAGAGAGCCGTTCCGCTGCCCATCG 522
XX
XX 81 CysArgHisGlnValValLeuAspArgHisGlyValTyrgLysLeuGlnArgAsnLeuLeu 100
XX 523 TGCAGGCATGAGGTTGTCTGGACAGACGGTCTTACGGCCGTGCACGAAACCTGTCTA 582
XX
XX 101 ValGlnAsnIleLeuAspIleTyrgLysGlnGlnSerSerArgProLeuHisAlaLysAla 120
XX 583 GTGGAGAACATTTATGCACATTTACAGCAGAGATCATCCAGGCGCTGCACCTCAAGGCT 642
XX
XX 121 GlnGlnHisLeuMetCysGlnGlnHisGlnAspGlnLysIleAsnIleTyrgCysLeuSer 140
XX 643 GAGCAGACACTCTATGTCGAGAGCATGAGAGAGAGAGAGATATATTATCTGCTGAGC 702
XX
XX 141 CysGlnValProThrCysSerLeuCysLysValPheGlnAlaHisLysAspCysGlnVal 160
XX 703 TGTGAGTGCCTCCACTGTCTCTCTGTGCAAGGTCTTCGTTGCCCAAGAGACTGTGAGTG 762

```

```

XX 161 AlaProLeuProThrIleTyrgLysArgGlnLysSerGlnLeuSerAspGlyIleAlaMet 180
XX 763 GCCCACTGCCACCATTATACAAAGCCAGAGAGTGAAGCTACAGCATGCCGATG 822
XX
XX 181 LeuValAlaGlnLysAsnAspArgValGlnAlaValIleThrGlnMetGlnLysValCysGln 200
XX 823 CTGATGTCAGGACATGACCGCGTGCACAGCATGATCACACAGATGAGAGAGTGTCCAG 882
XX
XX 201 ThrIleGlnAspAsnSerArgArgGlnLysGlnLeuLeuAsnGlnArgPheGlnThrLeu 220
XX 883 ACATATCAGGACATATACCGGAGGACAGAGCATGTTTAAACAGATTAGAGCCCTG 942
XX
XX 221 CysAlaValLeuGlnGlnArgLysGlnLysLeuLeuGlnAlaLeuAlaArgGlnGln 240
XX 943 TCGGCACTGCTGGAGAGAGCGCAAGGTTAGTCTGCTCCAGCGCTGCCCGGAGCAAGAG 1002
XX
XX 241 GlnLysLeuGlnArgValArgGlyLeuIleArgGlnTyrgLysPheHisLeuGlnLysSer 260
XX 1003 GAGAGCTGCGAGCGCGTCCGCGGCTCATCCGTCAGTATGGCGACCACTGGAGGCTCC 1062
XX
XX 261 SerLysLeuValGlnSerAlaIleGlnSerMetGlnGlnProGlnMetAlaLeuTyrgLeu 280
XX 1063 TCTAAGCTGTGGAGTGTGCTCCATCCAGTCATGAGAGAGCCCAATGCGCTGTATCTC 1122
XX
XX 281 GlnGlnAlaLysGlnLeuIleAsnLysValGlyAlaMetSerLysValGlnLeuAlaGly 300
XX 1123 CAGCAGGCCAAGAGAGCTGATCAATTAAGTCCGAGGCGCATGTCAGAGTGCAGCGAGGG 1182
XX
XX 301 ArgProGlnProGlyTyrgLysSerMetGlnGlnPheSerValSerValGlnHisValAla 320
XX 1183 CGGCGGAGGACGAGCTATGAGAGCATGAGCATGAGCATTCACCGTAAAGGTGAGAGCACGCGGCC 1242
XX
XX 321 GlnMetLeuArgThrIleAspPheGlnProGlyAlaAlaGlyAspGlnLysAspAsp 340
XX 1243 GAAATGCTGGGACATCTGACTTCAGGCGAGCGCTCCGGGGAGAGAGAG-----GAG 1296
XX
XX 341 MetAlaLeuAspGlyGlnGlnGlnLysAlaGlnLysGlnGlnLysGlnLysValPro 360
XX 1297 CTGCGCCCAAGAGAGAGAGAGAGAGGAGCGCGCGGAGAGAGAGAGAGAGAGAGAGAGAG 1356
XX
XX 361 GlnGly 362
XX
XX 1357 TAAAGC 1362
XX
XX RESULT 4
XX ID AAA72433 standard; cDNA; 1349 BP.
XX AC AAA72433;
XX XX
XX 19-DEC-2000 (first entry)
XX
XX Human nucleic acid-binding protein NuABP-52 cDNA.
XX
XX Human nucleic acid-binding protein; NuABP; agonist; antagonist; EST;
XX expressed sequence tag; drug screening; recombinant expression; antibody;
XX reproductive disorder; infertility; immunological disorder;
XX neurological disorder; cell proliferative disorder; cancer; tumour; ss.
XX
XX Homo sapiens.
XX
XX MO200044900-A2.
XX
XX 03-AUG-2000.
XX
XX 28-JAN-2000; 2000MO-US02237.
XX PF
XX 29-JAN-1999; 99US-0117904.
XX PR
XX 29-JAN-1999; 99US-0117905.
XX
XX (INCY-) INCYTE PHARM INC.
XX

```

Tang YT, Lai P, Hillman JL, Yue H, Azimzal Y, Lu AMD, Baughm MR, Tran B, Shih LL, Au-Young JL;  
P-PSDB; AAB21048.  
MPI; 2000-499332/44.  
Novel nucleic acid binding proteins, used to identify agonists and antagonists of them, for the treatment of reproductive, immunological, neurological and cell proliferative disorders including cancer -  
Claim 4; Page 178-179; 180pp; English.

Sequences AAT72382/472436 represent cDNAs encoding novel human nucleic acid-binding proteins (NuABPs; BAB20997-B21051). These cDNAs were produced by extension from an appropriate EST (expressed sequence tag) using primers designed using the EST. The invention also relates to expression constructs, host cells and transgenic organisms comprising a human NuABP nucleic acid, recombinant production of the human NuABPs, and antibodies against the human NuABPs, and also to methods of screening modulators of human NuABP activity or expression. The human NuABPs, and their agonists and antagonists are used to treat diseases associated with overexpression or underexpression of functional NuABPs. Human NuABP proteins and nucleotides, and NuABP agonists and antagonists can be used to diagnose, treat and prevent reproductive, immunological, neurological and cell proliferative disorders. Reproductive disorders that may be treated using compositions of the invention include infertility, endometriosis, disruptions of the menstrual cycle and disruptions of spermatogenesis. Immunological disorders that may be treated include AIDS, allergies, and autoimmune disorders such as multiple sclerosis, rheumatoid arthritis, diabetes and systemic lupus erythematosus. Neurological disorders that may be treated include epilepsy, neurodegenerative conditions such as Alzheimer's disease and Parkinson's disease, prion diseases such as Creutzfeldt-Jakob disease, and mental disorders such as schizophrenia. Cell proliferative disorders that may be treated include a wide variety of cancers, and also arteriosclerosis, atherosclerosis, cirrhosis and psoriasis.

Sequence 1349 BP; 328 A; 372 C; 428 G; 221 T; 0 other;

Alignment Scores:	
- No.:	4,366-160
- Length:	1349
- Mismatches:	1725,00
- Conservative:	337
- Local Similarity:	96,168
- Mismatches:	14
- Local Similarity:	92,338
- Mismatches:	8
- Local Similarity:	90,228
- Mismatches:	6
- Local Similarity:	21
- Mismatches:	2

US-09-908-988B-2 (1-366) x AAA72433 (1-1349)

QY	1	MeksniphetrValGlyPhelYsProLeuLeuGlyAspAlaHisAsnAspAsnLeu	20
Db	129	ATGAACTTCAAGTGGGTTTCAAGCCGCTCTAGGGGATGCACACAGATGCACACCTG	188
QY	21	GIuLySGlnLeuIleCysProIleCysLeuGIuMeCPheSerLysProValIleLeu	40
Db	189	GAGAAAGAGCTCACTCTCCCATCTGCTGGTAATCTTCCAAACAGAGGTATCTTG	248
QY	41	ProCysGlnHisAsnLeuCysArgLysCysAlaAsnArgValPheGlnIleSerAsnPro	60
Db	249	CCCTGCCAACACAAACGTGGCGGGAATGTGCCAACAGACGCTTCCAGGCGCTGGAATCCT	308
QY	61	LeuTProIleSerArgGlySerThrThrValSerSerGlyLysArgPheArgCysProSer	80
Db	309	CTATGGCAGTCCGGGGGCTCCACACAGTGTCTTCCAGGAAGCGGTTCCGCTCCCATCG	368
QY	81	CysArgHisGlnValValLeuAspArgHisGlyValTyrGlyLeuGlnArgAsnLeu	100
Db	369	TGCAGGCATGAGGTGGTTCGCTGGAAGACACAGTGTCTACGGGCTGGACGCAAAAGTGTCTA	428
QY	101	ValGIuAsnIleLeuAspIleTyrLysGlnGlnSerSerArgProLeuHisAlaLysAla	120
Db	429	GTTGGAGAACATTATCGACATTATTCAGACAGAGATCATCAAGCGCGCTGCATGAAGGCT	488

Oy	121	GAGCAGCACTTCATGTGGAGAGCATGAAAGACGAATTAATTACTGCTGAGC	548
Oy	141	CysglValProThrCysSerLeucCysIleValPheGlyAlaHisIleAspCysglVal	160
Db	549	TGTGAGTGCCACCACCTGCCTCTCTCGCAAGGCTCTCGGTGCCAACAGACTGTGAGGTG	608
Oy	161	AlaProLeuProThrlIeTyrIleYsrArgGlnIlysserGluLeuSerAspGlyIleAlamet	180
Db	609	GCCCCACAGCCACCATTTTACAAAAGCCAGAAGAGTAGCTCACGAGAGGATGCGCATG	668
Oy	181	LeuValAlaGlyAsnAspArgValGlnAlaValIleTrpArgImetGluIvalCysGln	200
Db	669	CTGGTGGCAGGCAATAGACCGCGGTGCACAACAGTATGACAGATGAGAGAGTGTGCCAG	728
Oy	201	ThrIleGluAspAsnSerArgArGlnIlysglnLeuIleuAsnglnIaryPheGluThrIeu	220
Db	729	ACTATCGAGGACAAATAGCCGGAGGAGAGACAGTTGTMAACACAGAGGTTTTGAGAGCCTG	788
Oy	221	CysAlaValIleGluGlnArgIylsGlyGlnIueuenglnAlaleuAlaArgGluGlnGlu	240
Db	789	TGCCAGAGTCTGGAGAGACGCAAGGTAGCTGCTGCAGCGCGCTGGCCCCGGAGCAAGG	848
Oy	241	GluIylsleugInaryValAryGlyLeuIleargGlnTyrgIyaSPHSLeugluclyser	260
Db	849	GAGAACTGCGACGCGCTCCGCGGCTCATCCGTAGTATGGCGACACACCTGGAGGCTCC	908
Oy	261	SerIylsleuValGluSerAlaIleGlnSermetGluGluProGlnmeAlaIleuTyrlieu	280
Db	909	TCTATAGCTGTGGATGATCTGCGCATCCAGTCCATGGAGAAGCCCAAATGGCCCTGTATCTC	968
Oy	281	GlnGlnAlaIylsGluIeuIleasnIylsValGlyAlameSertIylsValGluIleuAlagly	300
Db	969	CAGCAGCGCCAAGAGACTGATCAATPAAGTTCGGGGCCATGCTGGAAAGTGTGAGCTGGCAAGG	1028
Oy	301	ArgProGluProGlyTYrGluSermetcLuginPheserValSerValGluHisValala	320
Db	1029	CGCGCGGAGCCAGGCTTATGAGAGCATGAGCAATTCACCGCTMAAGGTGGAGACAGCTGGCC	1088
Oy	321	GluMetIleArcgThrIleaspPheGlnProGlyAlaIala-GlyaspGluGluAspAspas	340
Db	1089	GAATGCTGCGGACCATGACTCTCCAGCCAGGCGCTTCGGGGGAGGAAAGG-----GA	1142
Oy	340	pMeAlaIeuaspgly-gluGluGlynalaglyLeuGluGluGluIaryGruIeuaspyalp	360
Db	1143	GGTGGCCCCCAAGCAGGAAGAGGGCAACCGGG------GCCGGAAGAAAAAAGGCC	1196
Oy	360	roGluGlySer 363	
Db	1197	CGGATGGGGCC 1207	
<b>RESULT 5</b>			
ABN85313	ID	ABN85313 standard; cDNA, 2040 BP.	
XX	AC	ABN85313;	
XX	DT	30-SEP-2002 (first entry)	
DE	XX	Human cytoskeleton-associated protein, CSAP-4, coding sequence.	
KW	XX	Human: cytoskeleton-associated protein; CSNP; CSNP-4;	
KW	XX	cell proliferative disorder; viral infection; neurological disorder;	
KW	XX	transgenic animal; antihypertensive; antisporadic; antiinflammatory	
KW	XX	vincidne; anticonvulsant; vasotropic; cerebroprotective; motropic;	
KW	XX	neuroprotective; cyostatic; gene therapy; gene; ss.	
OS	XX	Homo sapiens.	
TH	Key	Location/Qualifiers	
CT	CDS	291..1271	
CT		/tag= a	



/product= "CSAP-4"

FT xx MO200253719-A2.  
PN xx 11-JUL-2002.  
PD xx 04-JAN-2002; 2002WO-US00178.  
PE xx  
PR 04-JAN-2001; 2001US-260085P.  
PR 13-FEB-2001; 2001US-268554P.  
PR 14-FEB-2001; 2001US-26911P.  
PR 23-FEB-2001; 2001US-271211P.  
XX xx  
PA (INCY-) INCYTE GENOMICS INC.  
XX  
PI Lu DAM, Baughn MR, Yao MG, Ding L, Honchell CD, Yue H, Tang YT;  
PI Warren BA, Duggan BM, Xu Y, Walla NK, Griffin JA, Stewart EA;  
PI Gandhi AR, Khan FA, Thangaveju K, Ison CH, Azimzai Y, Hafalia AJA;  
PI Gietzen KU, Lal PG, Sanjanwala MM, Elliott VS;  
XX  
DR WPI: 2002-583611/62.  
XX P-PSDB; ABB83475.  
XX  
XX Novel isolated human cytoskeleton-associated protein for diagnosing,  
PT treating or preventing atherosclerosis, psoriasis, leukemia, epilepsy,  
PT ischemic cerebrovascular disease, cerebral neoplasms and Alzheimer's  
PT disease -  
XX  
XX Claim 5, Page 154-155; 167pp; English.  
XX  
XX The present sequence is the coding sequence for a human  
CC cytoskeleton-associated protein (CSAP). The CSAP and its coding sequence  
CC are useful in the diagnosis, treatment and prevention of a cell  
CC proliferative disorder such as actinic keratosis, atherosclerosis,  
CC psoriasis, primary thrombocythemia, leukemia; a viral infection such as  
CC those caused by adenoviruses (acute respiratory disease, pneumonia),  
CC arenaviruses (lymphocytic choriomeningitis); and a neurological disorder  
CC such as epilepsy, ischemic cerebrovascular disease, stroke, cerebral  
CC neoplasms, Alzheimer's disease, Pick's disease, Huntington's disease or  
CC amyotrophic lateral sclerosis. The CSAP coding sequence is also useful  
CC for creating knock out or knock in humanised animals or transgenic  
CC animals to model human diseases.  
XX  
XX Sequence 2040 BP; 465 A; 584 C; 650 G; 341 T; 0 other;  
XX  
XX  
XX Alignment Scores:  
Pred. No.: 1,94e-147 Length: 2040  
Score: 1599.00 Matches: 311  
Percent Similarity: 88.67% Conservative: 10  
Best Local Similarity: 85.91% Mismatches: 7  
Query Match: 83.63% Indels: 34  
Gaps: 2  
US-09-908-988b-2 (1-366) x ABBN5313 (1-2040)  
QY 1 MetAspHeThrValGlyPheLysProLeuLeuGlyAspAlaHisAsnMetAspAsnLeu 20  
DB 291 ATGAACTTACAGTGGGTTTCAAGCCCTGCTAGGGAGTGCACACAGATGACAACTCG 350  
QY 21 GluLysGlnLeuIleCysProIleCysLeuGlnMetPheSerLysProValIleLeu 40  
DB 351 GAGAAACACATCTGCCCCATCTGCTGAGATGTTCTCCAAACAGTGGTATCTCG 410  
QY 41 ProCysGlnHisAsnLysArgLysCysAlaAsnAspValPheGlnAlaSerAsnPro 60  
DB 411 CCCGCGCAACACACACTGTCGCGCAAAATGTCACACAGACGCTTCCAGGCCCTCGAATCCT 470  
QY 61 LeuTrpGlnSerArgLysSerThrThrValSerSerGlyGlyArgPheArgCysProSer 80  
DB 471 CTAATGCAATCCCGGGGCTCCACACACTGTGTCTTCAGAGGCCGCTTCCGCTGCCCATCG 530  
QY 81 CysArgHisGlnValValLeuAspArgHisGlyValTyrGlyLeuGlnArgAsnLeuLeu 100  
|||||

DB 531 TGCAGCATGAGGTTCCTCTGACAGACACGGTGTCTAGCGGCTGCAGCAAACTGCTA 590  
QY 101 ValGluAsnIleIleAspIleTyrLysGlnGluSerSerArgProLeuHisAlaLysAla 120  
DB 591 GTGAGACATATTATCGACATTATACAGAGAGTATCATCGAGCCGCTGCATCCAAAGGCT 650  
QY 121 GlnGlnHisLeuMetCysGlnGlnHisGlnAspGlnLysIleAsnIleTyrCysLeuSer 140  
DB 651 GAGAGACACCTCATGTGCGAGAGCATGAAGAAGAAATCAATATTACTGCTCCGAGC 710  
QY 141 CysGlnValProThrCysSerLeuCysLysValPheGlyAlaHisLysAspCysGlnVal 160  
DB 711 TGTGAGGTGCCCGACCTGCTCTCTGCAAGGCTTCGCGCCCAAGAGACTGTGAGGTG 770  
QY 161 AlaProLeuProThrIleTyrLysArgGlnLysSerGluLeuSerAspLysIleAlaMet 180  
DB 771 GCCCACTGCCCCACCATTTTACAAACGCCAGAG----- 803  
QY 181 LeuValAlaGlyAsnAspArgValGlnAlaValIleThrGlnMetGluValCysGln 200  
DB 803 ----- 803  
QY 201 ThrIleGluAspAsnSerArgArgGlnLysGlnLeuLeuAsnGlnArgPheGlnThrLeu 220  
DB 804 -----GACATATAGCCGAGAGCAGAAAGCTTTGTTAAACAGAGTTTGAGACCTG 854  
QY 221 CysAlaValLeuGlnGluArgLysGlnLeuLeuGlnAlaLeuAlaArgGlnGln 240  
DB 855 TGGCAATGCTGGAGAGACGCAAGGGTAGCTGTCAGAGGCTGGCCGGGAGCAAGAG 914  
QY 241 GluLysLeuGlnArgValArgGlyLeuIleArgGlnTyrGlyAspHisLeuGlnLysSer 260  
DB 915 GAGAGAGTGCAGCGCGCTCCGGGCTCATCCGTAGTATGGGAGCACCTGGAGGCTCC 974  
QY 261 SerLysLeuValGluSerAlaIleGlnSerMetGluLysProGlnMetAlaLeuTyrLeu 280  
DB 975 TCTAACCTGGGTGGAGTCTGCCATCCAGTCCATGGAAGACCAAAATGCGCTGTATCTC 1034  
QY 281 GlnGlnAlaLysGlnLeuIleAsnLysValGlyAlaMetSerLysValGlnLeuAlaGly 300  
DB 1035 CAGAGGCCCAAGAGAGCTGATCAATATAGGTCCGGGCAATGTCAGAGGTGAGCTGGCAGGG 1094  
QY 301 ArgProGluProGlyTyrGluSerMetGluGlnPheSerValSerValGlnHisValAla 320  
DB 1095 CGGCGGAGGACGAGGCTATGAGAGCATGAGACATTCACCGTAAAGGTGAGAGCAGTGGCC 1154  
QY 321 GlnMetLeuArgThrIleAspPheGlnProGlyAlaAlaGlyAspGlnLysAspAsp 340  
DB 1155 GAATGCTGCGGACCATCTGACCTTCAGCCAGCGCTTCCGGGGAGGAAAGAG-----GAG 1208  
QY 341 MetAlaLeuAspGlyGlnGlnGlyAsnAlaGlyLeuGlnGlnGlnArgLeuAspValPro 360  
DB 1209 GTGCCCCAGACGAGAGAGAGGAGCGCGGGGCCGGAGGAGAGCGCGGAGTGGGCTT 1268  
QY 361 GluGly 362  
DB 1269 TAAGGC 1274  
RESULT 6  
AAAF27653  
ID AAF27653 standard; DNA: 1500 BP.  
XX  
AC AAF27653:  
XX  
XX 30-MAR-2001 (first entry)  
DE DNA encoding human transcriptional regulator protein #22.  
XX  
KW Human; transcriptional regulator protein; TXREG; ds.  
XX  
OS Homo sapiens.  
XX  
PN WO200078954-A2.



```

XX 28-DEC-2000.
PD
PF 15-JUN-2000; 2000WO-US16766.
XX
PR 18-JUN-1999; 99US-0140109.
XX
PA (INCY-) INCYTE GENOMICS INC.
XX
PI Lal P, Yue H, Tang YT, Baughn MR, Azimzai Y, Tran B;
DR WPI: 2001-041425/05.
XX
PT Isolated polypeptide with a human transcriptional regulator protein
PT sequence is useful for the diagnosis, prevention and treatment of
PT disorders associated with the immune, reproductive and cardiovascular
PT systems.
XX
XX Claim 5; Page 135-136; 142pp; English.
XX
CC The present invention relates to human transcriptional regulator
CC protein (TXREG) sequences. The antagonist and an agonist of the proteins
CC of the invention are used to treat disorders associated with decreased
CC or increased expression or activity of TXREG.
XX
SQ Sequence 1500 BP; 401 A; 378 C; 439 G; 282 T; 0 other;
XX
Alignment Scores:
Pred. No.: 2,11e-114 Length: 1500
Score: 1261.50 Matches: 251
Percent Similarity: 89.04% Conservative: 17
Best Local Similarity: 83.39% Mismatches: 27
Query Match: 65.98% Indels: 6
DB: 22 Gaps: 4
XX
US-09-908-988b-2 (1-366) x AAF27653 (1-1500)
QY 1 MetanpherhrvalglphelysProleuLeuGlyAspAlaHisAsnMetAspAsnLeu 20
DB 291 ATGAACCTTCACAGTGGGTTTCAAGCCGCTGCTAGGGAGATGCACACAGCATGCACACCTG 350
QY 21 GluysGlnLeuIleCysProIleCysIleuGluMetPheSerIysProValAlIleLeu 40
DB 351 GAGAGAGCGCTCATCTGCCCCATCTGCCCTGGAGATGTTCTCCAAACACAGTGGATCTG 410
QY 41 ProCysGlnHisAsnLeuCysArgIysCysAlaAsnAspValPheGlnIleSerAsnPro 60
DB 411 CCCGCCACACACACCTTGGCCGCAATGTGCCAAGAGAGCTTCCAGGCTTGGAATCCT 470
QY 61 LeuTrpGlnSerArgIySerThrThrValSerSerGlyIArgPheArgCysProSer 80
DB 471 CTATGGCAGTCCCGGGGCTCCACCACTGTGTCTCAGAGGCGGTTCCGCTGCCATCG 530
QY 81 CysArgHisGluValAlValLeuAspArgHisGlyValTyrGlyLeuGlnArgAsnLeu 100
DB 531 TGCAGGCGCTGAGGTGTCTCTGACAGACAGCGGTGTCTACGGCTTCAGGAAACCTGCTA 590
QY 101 ValGluAsnIleIleAspIleTyrIysGlnGlnUserSerArgProLeuHisAlaIysAla 120
DB 591 GTGAGAAACATTTATGACATTTTACAGCAGAGGATCATCCAGGCCCTGCACCTCAAGGCT 650
QY 121 GluGlnHisLeuMetCysGluGlnHisGlnAspGluIysIleAsnIleTyrCysLeuSer 140
DB 651 GAGCAGACACCTCATGTGTGAGAGCATGAGAAAGAAAGATCAATTTTACGCTCAGAGGCT 710
QY 141 CysGluValProThrCysSerLeuCysIysValPheGlyAlaHisIysAspCysGluVal 160
DB 711 TGTAGAGTGCCCACTGCTCTCTGTGCAAGGTCTTCGGTCCCAAGAGAGCTGTGAGGTG 770
QY 161 AlaProLeuProThrIleTyrIysArgIlnIysSerGluLeuSerAspGlyIleAlaMet 180
DB 771 GCCCACTGCCCACTTTTACAAAGCCAGAAAGAGTGAAGCTCAGCATGATGCGGATG 830

```

```

QY 181 LeuValAlaGlyAsnAspArgValGlnAlaValIleThrGlnMetGlnGluValCysGln 200
DB 831 CTGTGGCAGGCAATGACCGCGCTGCAAGCATGATCACAAGATGAGAGAGGTGTCCAG 890
QY 201 ThrIleGluAspAsnSerArgArgIlnIysGlnLeuLeuAsnGlnArgPheGlnThrLeu 220
DB 891 ACTATCGAGGACAATFAGCCGGGAGCAGAAAGCAGTGTGTTAACACAGAGGTTGAGACCTG 950
QY 221 CysAlaValIleuGluArgIysGlyGluLeuLeuGlnAlaIleuAlaArgGluGln-GI 240
DB 951 TGCAGAGTCTGGAGAGACCGCAGCGTGTGAGTGTGACAGGCGCTGCGCGGAAACAAGCA 1010
QY 240 uGluIysLeuGlnArg---ValArgGlyLeuIleArg---GlnTyrGlyAspHisLeuI 258
DB 1011 GGACAGAGCTTCAAGCGGATCCAGCGAGCTCATTTCCGGTCAATATGCGAAGCAACTT 1070
QY 258 uGlySerSerIysLeuValGluSerAlaIle-----GlnSerMetGluIlnProGlnMe 276
DB 1071 GGGAAGGCCCTCCCTGCTAAAGCGCTGTTGAGAGGCTCACTGTCCAAATCCCAAGTTT 1130
QY 276 tAlaLeuTyrIleuGlnGlnAlaIysGluLeuIleAsnIysValGlyAlaMetSerIysVa 296
DB 1131 CCACATTTGGATATACAGAGC---CCAACAGAAATTTGTGTGCGGACACTTAGTAAAT 1187
QY 296 1 296
DB 1188 G 1188
XX
RESULT 7
ABA99062
ID ABA99062 standard; DNA; 2590 BP.
XX
AC ABA99062;
XX
DE 15-JUL-2002 (first entry)
XX
DE Murine muscle ring finger protein 2 (MURF-2) coding sequence.
XX
KW Muscle ring finger; MURF-2; mouse; cardiac; microtubule;
KW intermediate filament; striated muscle; cardiac hypertrophy;
KW heart disease; gene; ds.
XX
OS Mus musculus.
XX
FH Key Location/Qualifiers
FT CDS 80..1717
FT FT /*tag= a
FT FT /product= "MURF-2"
XX
MO200206318-A2.
XX
PD 24-JAN-2002.
XX
PF 18-JUL-2001; 2001WO-US22896.
XX
PR 18-JUL-2000; 2000US-219020P.
XX
PA (TEXA ) UNIV TEXAS SYSTEM.
XX
PI Olson EN, Spencer JA;
XX
DR WPI: 2002-241506/29.
XX
DR P-RSDB; ABB08276.
XX
PT Novel muscle ring finger protein useful for drug screening, and for
PT diagnosing and treating diseases, particularly cardiomyopathies
XX
PS Claim 4; Page 126-129; 134pp; English.
XX
CC The sequence encodes murine muscle ring finger protein 2 (MURF-2). The
CC invention relates to a purified muscle ring finger (MURF) protein,
CC selected from MURF-1, MURF-2 and MURF-3. The MURF proteins of the
CC invention are involved in microtubule and intermediate filament

```

CC stabilisation of striated muscle cells and have cardiant activity.  
 CC The MURF proteins are useful for screening a candidate substance  
 CC for MURF protein binding activity, in a cell, cell-free system or  
 CC in vivo, and its effect on interaction of MURF with microtubules,  
 CC homodimerisation of MURF, MURF-1, MURF-2 or MURF-3 stabilisation  
 CC of microtubules, interaction of MURF with intermediate filaments,  
 CC e.g. desmin, vimentin and cytokeratin, and heterodimerisation of MURF.  
 CC The screened compounds are useful for treating and preventing of  
 CC cardiac hypertrophy and heart diseases. MURF proteins are useful as  
 CC antigens to immunise animals for the production of antibodies.  
 XX  
 SQ Sequence 2590 BP; 703 A; 627 C; 625 G; 635 T; 0 other;  
 Alignment Scores:  
 Pred. No.: 2.78e-104 Length: 2590  
 Score: 1162.00 Matches: 226  
 Percent Similarity: 76.99% Conservative: 55  
 Best Local Similarity: 61.92% Mismatches: 76  
 Query Match: 60.77% Indels: 8  
 Gaps: 4  
 DB: 24  
 US-09-908-988b-2 (1-366) x ABA99062 (1-2590)  
 QY 1 MetasphethrValGlyPheLysProLeuLeuGlyAspAlaHisAsnMetAspAsnLeu 20  
 DB 80 ATGAGCACTTCTCTGAATTACAGTCTTCTCCAAAGAGCAGCAGACCATGATACCTG 139  
 QY 21 GlnGlyGlnLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 40  
 DB 140 GAAAGCACTGATGTCCTCCATCTCCAGATGTCCTCCAGATGTCCTCCAGATGTCCTCC 199  
 QY 41 ProGlyGlnHisAsnLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 60  
 DB 200 CTTTCCAGCAGACACCTGTCGAGAAATGTCCTCCAGATGTCCTCCAGATGTCCTCCAG 259  
 QY 61 LeuTrpGlnSerArgGlySerThrThrValSerSerGlyGlyArgPheArgCysProSer 80  
 DB 260 TACTTACCCACAGAGAGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 319  
 QY 81 CysArgHisGlnValValLeuAspArgHisGlyValValValValValValValValVal 100  
 DB 320 TGCACACATGAGTGGTGTAGACACAGATGGGTGTATGACATGACAGAGAACCTGCTC 379  
 QY 101 ValGluAsnAlaLeuAspLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 120  
 DB 380 GTGGAACATTAATGATGATCTACAGCAGAAATCCACAGGCCA---GAAAAAATG 436  
 QY 121 GlnGlnHisLeuMetCysGlnGlnHisGlnAspGluLysIleAsnIleTrpCysLeuSer 140  
 DB 437 GACCAAGCCC---ATGCTGTAAGAGCATGAAGAGAGCATCAACATCTATTGCTGTAAC 493  
 QY 141 CysGlnValProThrCysSerLeuCysLysValPheGlyIleHisLysAspCysGlnVal 160  
 DB 494 TGTGAAGTCCACCATGCTCTGTCCTGCAAGGTTTGGCCCTTAAGGATGTCAGAGTG 553  
 QY 161 AlaProLeuProThrIleLeuTrpArgGlnLysSerGluLeuSerAspGlyIleAlaMet 180  
 DB 554 GCTCCCTGACATCATCTGTTCCAGAGCAGACATCAGAGCTCATGATGATGCTGTA 613  
 QY 181 LeuValAlaGlyAsnAspArgValGlnAlaValIleThrGlnMetGlnGlnValCysGln 200  
 DB 614 CTGTGTGAGAACAGATAGAGTCCAGAGGTGATCAGCAGCTGAGAGACACCTGTAA 673  
 QY 201 ThrIleGlnAspAsnSerArgArgGlnLysGlnLeuLeuAsnGlnIleArgPheGluThrLeu 220  
 DB 674 ACTAATGAGAGAGTCTGCTGCAAGAAAGCAGACCTGTGTGAGAAATTTGACCTCA 733  
 QY 221 CysAlaValLeuGlnGlnLysArgLysGlnLeuLeuGlnAlaLeuAlaArgGlnGlnLys 240  
 DB 734 TACGCATCTGAG 793  
 QY 241 GlnLysLeuGlnArgValArgGlyLeuIleArgGlnIleArgLysAspHisLeuGlnLys 260

DB 794 GAGAACTGGAACTGTCGCACTTATCAGAGAGTATTCGATTCACCTGGAGACGTA 853  
 QY 261 SerLysLeuValGlnSerAlaIleGlnSerMetGlnGlnProGlnMetAlaLeuTrpLeu 280  
 DB 854 TCCAACTGGTGGAGTCAGGAATCCAGTTGATGATGAGCCGAAATGGCAGTATTTCTG 913  
 QY 281 GlnGlnAlaLeuGlnLeuLeuIleAsnLysValGlyAlaMetSerLysValGlnLeuAlaGly 300  
 DB 914 CAGATGCCAAGACCTGTTGCAAAAGATGCTGGAAGCATCAAAAGCCTTCAGATGAG 973  
 QY 301 ArgProGlnProGlyTrpGlnSerMetGlnGlnPheSerValSerValGlnHisValAla 320  
 DB 974 AAACATGACCAAGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1033  
 QY 321 GlnMetLeuArgThrIleAspPheGlnProGlyAlaAlaGlyAspGlnGlnAspAsp 340  
 DB 1034 AAATTAATCCGTAATATGACTTT-----TCTAGAGAGAGAGAGAGAGAGAGAGAT 1084  
 QY 341 MetAlaLeuAspGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 360  
 DB 1085 GCAGAGAAATAGATGAG 1135  
 QY 361 GlnGlySerGlyLeu 365  
 DB 1136 GAGGCGAAGAAATGTT 1150  
 RESULT 8  
 AAS25855  
 ID AAS25855 standard; cDNA, 1762 BP.  
 XX  
 AC AAS25855;  
 XX  
 DT 07-NOV-2001 (first entry)  
 XX  
 DE Human cDNA encoding a novel secreted protein, Seq ID 34.  
 XX  
 KW Human; immunosuppressive; antiarthritic; ss; antirheumatic;  
 KW cytosolic; cardiant; vasotropic; cerebroprotective; nootropic;  
 KW neuroprotective; antibacterial; virucide; fungicide; opthalmlological;  
 KW vulnary; secreted protein; rheumatoid arthritis;  
 KW hyperproliferative disorder; cardiovascular disorder; cardiac arrest;  
 KW cerebrovascular disorder; cerebral ischemia; angiogenesis;  
 KW nervous system disorder; Alzheimer's disease; infection; ocular disorder;  
 KW corneal infection; wound healing; epithelial cell proliferation;  
 KW skin aging; food additive; preservative; antiproliferative.  
 OS Homo sapiens.  
 XX  
 PN WO200155322-A2.  
 XX  
 PD 02-AUG-2001.  
 XX  
 PF 17-JAN-2001; 2001MO-US01341.  
 XX  
 PR 31-JAN-2000; 2000US-0179065.  
 PR 04-FEB-2000; 2000US-0180628.  
 PR 24-FEB-2000; 2000US-0184664.  
 PR 02-MAR-2000; 2000US-0186350.  
 PR 16-MAR-2000; 2000US-0189874.  
 PR 17-MAR-2000; 2000US-0190076.  
 PR 18-APR-2000; 2000US-0198123.  
 PR 19-MAY-2000; 2000US-0205515.  
 PR 07-JUN-2000; 2000US-0209467.  
 PR 28-JUN-2000; 2000US-0214886.  
 PR 30-JUN-2000; 2000US-0215135.  
 PR 07-JUL-2000; 2000US-0216647.  
 PR 07-JUL-2000; 2000US-0216880.  
 PR 11-JUL-2000; 2000US-0217487.  
 PR 14-JUL-2000; 2000US-0218290.  
 PR 26-JUL-2000; 2000US-0220963.  
 PR 26-JUL-2000; 2000US-0220964.  
 PR 14-AUG-2000; 2000US-0224518.





CC of microtubules, interaction of MURF with intermediate filaments,  
 CC e.g. desmin, vimentin and cytokeratin, and heterodimerisation of MURF.  
 CC The screened compounds are useful for treating and preventing  
 CC cardiac hypertrophy and heart diseases. MURF proteins are useful as  
 CC antigens to immunise animals for the production of antibodies.

XX Sequence 1597 BP: 405 A; 404 C; 438 G; 350 T; 0 other;

## Alignment Scores:

Pred. No.:	1.88e-99	Length:	1597
Score:	1110.00	Matches:	212
Percent Similarity:	78.53%	Conservative:	55
Best Local Similarity:	62.35%	Mismatches:	63
Query Match:	58.05%	Indels:	10
DB:	24	Gaps:	3

US-09-908-988B-2 (1-366) x ABA99063 (1-1597)

QY 4 ThrValAGlyPheLys-----ProLeuLeuGlyAspAlaHisAsn 16  
 Db 278 ACAGTCGCATTCACAAAGCAATATGATTTAAATCTAGCGCTCTCTGATGGAACGCT 337  
 QY 17 MetAspAsnLeuGluLysGlnLeuLeuIleCysProIleCysLeuGluMetPheSerLysPro 36  
 Db 338 ATGGAGAACTTGAGAGAGAGAGCTGATCTGCCCATCTGCGTGGAGATGTTTACCAGGCT 397  
 QY 37 ValValIleLeuProCysGlnHisAsnLeuCysArgLysCysAlaAsnAspValPheGln 56  
 Db 398 GTGGTCATCTCCCTCGCCACACACACACTCTGCGGAGATGTGCCACACACATCTCCAG 457  
 QY 57 AlaSerAsnProLeuTrpGlnSerArgLysSerThrValSerSerGlyLysArgPhe 76  
 Db 458 GCTGCGAATCCCTACTGAGCCACACCGCGGTGCTGCTGCTCATCTGTGAGAGCTGTTTC 517  
 QY 77 ArgCysProSerCysArgHisGlnValValLeuAspArgHisGlyValArgLysLeuGln 96  
 Db 518 CGTGCCTCCCTGCTGCCCATGATGATGATGACGGGACGCGGTGATGCGCTGAC 577  
 QY 97 ArgAsnLeuLeuValGluAsnIleIleAspIleTyrLysGlnLys--SerSerArgPro 115  
 Db 578 AGGAACCTGCTGTGGAAACATCATTTGACATCTACACAGAGAGTCCCATCTGCGCCC 637  
 QY 116 LeuHisAlaLysAlaGluGlnHisLeuMetCysGluGluHisGluAspGluLysIleAsn 135  
 Db 638 CTG-----CAGAAAGGACGACCCGATGTGCAGAGAACACGAGAACGAGAGATCAAC 691  
 QY 136 IleTyrCysLeuSerCysGluValProThrCysSerLeuCysLysValPheGlyAlaHis 155  
 Db 692 ATCTACTCTCTACGTGTGAGGTGCTACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 751  
 QY 156 LysAspCysGluValAlaProLeuProThrIleTyrLysArgGlnLysSerGlnLeuSer 175  
 Db 752 CAGGCGCTGTAGAGTGGCCCTTTGGCAAGCATCTTCCAGAGACAGAGATGAGCTGACGT 811  
 QY 176 AspGlyIleAlaMetLeuValAlaGlyAsnAspArgValGlnAlaValIleThrGlnMet 195  
 Db 812 AACTGCATCTCATCTGCTGCTGCGGGAACGACCGAGTGCACAGCATCTCTCAGCTG 871  
 QY 196 GluGluValCysGlnThrIleGluAspAsnSerArgArgGlnLysGlnLeuLeuAsnGln 215  
 Db 872 GAGGAGCTGCTGACAGGTGACCAAGAGATACCCACAGAGTGAAGAGAGGAGCTGAGCAG 931  
 QY 216 ArgPheGlnThrLeuCysAlaValLeuGluGlnArgLysGlyGluLeuLeuGlnAlaLeu 235  
 Db 932 AAGTTTGACACCTCTACGCGCATCTCGATGAGAAGAGAGAGAGTGTGCGACGGGATC 991  
 QY 236 AlaArgGluGlnGluGlnLysLeuGlnArgValArgGlyLeuIleArgGlnTyrGlyAsp 255  
 Db 992 ACGGAGGAGGAGAGAGAGAGAGCTGAGCTTCAATCAGAGCTGTGATCTCTCCAGTACAGGAG 1051  
 QY 256 HisLeuGluGlnLysSerSerLysLeuValGluSerAlaIleGlnSerMetGluGlnProGln 275  
 Db 1052 CAGCTGGAAAAAGTCCACCAAGCTTGTGGAGACCGGCGCATCAGTCCCTGGATGAGCCCGGA 1111

QY 276 MetAlaLeuTyrLeuGlnGlnAlaLysGluLeuIleAsnLysValGlyAlaMetSerLys 295  
 Db 1112 GGGGCTACTCTCTCTCTCAAGTCCCAAGAGAGCTCATCAAGAGATTTAGAGCTCCAG 1171  
 QY 296 ValGluLeuAlaGlyArgProGluProGlyTyrGlnSerMetGluGlnPheSerValSer 315  
 Db 1172 GCGTCCAGCTGGGGAGACAGACAGAGCTTTGAGAACATGATGACTTCTCTGAGC 1231  
 QY 316 ValGluHisValAlaGluMetLeuArgThrIleAspPheGlnProGlyAlaAlaGlyAsp 335  
 Db 1232 TTAGAACATATAGCAGAGAGCTTGGAGGCGCATTTGAGGACAGTAAAGAGATGTGAT 1291  
 RESULT 10  
 AAS42490  
 ID AAS42490 standard; cDNA; 2110 BP.  
 XX AAS42490;  
 AC AAS42490;  
 DT 18-DEC-2001 (first entry)  
 DE Human cDNA encoding an mdm1 protein, clone LG:247384.1:2000MAY19.  
 XX  
 KW Human; molecules for disease detection and treatment; mdm1; ss;  
 KW Antiarteriosclerotic; hepatotropic; antipsoriatic; cytoprotective;  
 KW immunosuppressive; antidiabetic; antihypertensive; neuroprotective;  
 KW osteopathic; antiarthritic; cell proliferative disorder;  
 KW arteriosclerosis; cirrhosis; psoriasis; cancer; adenocarcinoma;  
 KW leukemia; breast cancer; autoimmune disorder; AIDS;  
 KW acquired immunodeficiency syndrome; Addison's disease;  
 KW diabetes mellitus; asthma; multiple sclerosis; osteoarthritis.  
 XX  
 OS Homo sapiens.  
 XX  
 PN W0200162922-A2.  
 XX  
 PD 30-AUG-2001.  
 XX  
 PF 21-FEB-2001; 2001WO-US05896.  
 XX  
 PR 24-FEB-2000; 2000US-0185213.  
 PR 16-MAY-2000; 2000US-0205232.  
 PR 17-MAY-2000; 2000US-0205285.  
 PR 17-MAY-2000; 2000US-0205286.  
 PR 17-MAY-2000; 2000US-0205287.  
 PR 17-MAY-2000; 2000US-0205323.  
 PR 17-MAY-2000; 2000US-0205324.  
 PR  
 XX  
 PA (INCY-) INCYTE GENOMICS INC.  
 PA  
 PI Panzer SR, Spiro PA, Banville SC, Shah P, Chalup MS, Chang SC;  
 PI Chen A, D'Sa SA, Amshey S, Dahl CR, Dam TC, Daniels SE;  
 PI Dufour GE, Flores V, Fong WT, Greenawalt LB, Hillman JL, Jones AL;  
 PI Liu TF, Roseberry AM, Rosen BH, Russo FD, Stockdreher TK, Daffo A;  
 PI Wright RJ, Yap PE, Yu JY, Bradley DL, Bratcher SR, Chen W;  
 PI Cohen HJ, Hodgson DM, Lincoln SE, Jackson S;  
 XX  
 DR WPI: 2001-570631/64.  
 DR P-PSDB: AAU25438.  
 DR  
 PT New disease detection and treatment molecule polynucleotides and  
 PT polypeptides; useful for diagnosis and treatment of arteriosclerosis,  
 PT cirrhosis, psoriasis, cancer, autoimmune disorders, diabetes mellitus,  
 PT asthma and multiple sclerosis  
 PT  
 PS Claim 1; Page 121; 183pp; English.  
 PS  
 XX The invention relates to novel human molecules for disease  
 XX detection and treatment (mdm1 proteins) and the polynucleotides encoding  
 XX them. The MDm1 polynucleotides and polypeptides are useful for diagnostic  
 XX and therapeutic purposes e.g. to diagnose and treat cell proliferative  
 XX disorders (e.g. arteriosclerosis, cirrhosis and psoriasis) cancers (e.g.  
 XX adenocarcinoma, leukemia and breast cancer) autoimmune disorders



CC encodes is useful for the diagnosis and/or prognosis of one  
 CC or more types of cancer. The polynucleotide and polypeptide can be  
 CC used as nutritional sources or supplements and in the screening of  
 CC chemical compounds as potential drugs.

XX Sequence 1231 BP; 332 A; 305 C; 361 G; 233 T; 0 other;

## Alignment Scores:

Align. No.:	1-78e-98	Length:	1231
Score:	1098.50	Matches:	210
Percent Similarity:	78.00%	Conservative:	63
Best Local Similarity:	60.00%	Mismatches:	66
Query Match:	57.45%	Indels:	11
DB:	22	Gaps:	4

US-09-908-988B-2 (1-366) x AAH90037 (1-1231)

```

QY 10 LeuLeuGlyAspAlaHisAsnMetAspAsnLeuGluGlyGlnLeuIleCysProIleCys 29
    |||...|||...|||...|||...|||...|||...|||...|||...|||...|||...|||
Db 167 CTGATCCAGAGTGGAGATCCCATGTGAGAACTTGGAGAACAGCTGATCTGCTTACTGCG 226

QY 30 LeuGluMetPheSerIleProValAlaIleLeuProCysGlnHisAsnLeuCysArgLys 49
    |||...|||...|||...|||...|||...|||...|||...|||...|||...|||...|||
Db 227 CTGGAGATGTTTACCAGACCATGCTGATCTTGGCCGACAGACACACCTGCTCCGGAAG 286

QY 50 CysAlaAsnAspValPheGlnAlaSerAsnProLeuTrpGlnSerArgIleSerThr 69
    |||...|||...|||...|||...|||...|||...|||...|||...|||...|||...|||
Db 287 TGTGCATGACATCTCTCCAGGCTGCMAATCCCTACTGACACCGCGGGGACACTCAGTG 346

QY 70 ValSerSerGlyIleArgPheArgCysProSerCysArgHisGluValIleLeuAspArg 89
    |||...|||...|||...|||...|||...|||...|||...|||...|||...|||...|||
Db 347 TCCATGTCTGGAGCGGCTTCCGCTGCCCTCCACCTCCGACAGAGTGTATCATGATGCTG 406

QY 90 HisGlyValIleTrpGlyLeuGlnArgAsnLeuValGluAsnIleLeuAspIleTrpLys 109
    |||...|||...|||...|||...|||...|||...|||...|||...|||...|||...|||
Db 407 CACGGAGTGTACGGGCTCAGAGAACCTCTGCTGAGAAACATCTTCACATCTTACAAA 466

QY 110 GlnGlu---SerSerArgProLeuHisAlaLysAlaGluGlnHisLeuMetCysGluGlu 128
    |||...|||...|||...|||...|||...|||...|||...|||...|||...|||...|||
Db 467 CAGAGATGCTCCAGTCCGCGCCGCTG-----CAGAAAGGCGAGTCCACCCATGTCCAGAG 520

QY 129 HisGluAspGluLysIleAsnIleTrpCysLeuSerCysGluValAlaProLeuProThrIleTrpLys 148
    |||...|||...|||...|||...|||...|||...|||...|||...|||...|||...|||
Db 521 CACGAGATGAGAAAAATCAACATCTACTGCTCTCAGCTGTGAGAGTGCCTCCATCTCCATG 580

QY 149 CysLysValPheGlyAlaHisLysAspCysGluValAlaProLeuProThrIleTrpLys 168
    |||...|||...|||...|||...|||...|||...|||...|||...|||...|||...|||
Db 581 TGCAGAGTGTGGATGATCACAAGGCTCCGAGGTGGCCCATTTGCAGAGTGTCTTCCAG 640

QY 169 ArgGlnLysSerGluLeuSerAspGlyIleAlaMetLeuValAlaGluAsnAspArgVal 188
    |||...|||...|||...|||...|||...|||...|||...|||...|||...|||...|||
Db 641 GGACAAAAGATGAACTGAATACATGATCTTCATGCTGTGGCGGAGATGACCTGTG 700

QY 189 GlnAlaValIleThrGlnMetGlnGluValCysGlnThrIleGluAspAsnSerArgVal 208
    |||...|||...|||...|||...|||...|||...|||...|||...|||...|||...|||
Db 701 CAGACCATCATCTACCTCAGCTGAGGATTCCTCCGTGAGTGCACCAAGAGAACTTCAACAC 760

QY 209 GlnLysGlnLeuLeuAsnGlnArgPheGlnThrLeuCysAlaValLeuGluGlnArgLys 228
    |||...|||...|||...|||...|||...|||...|||...|||...|||...|||...|||
Db 761 GTAAAGGAGAGAGCTGAGCCAGAGTTTACACACGTTGTATGCTGATGAGAAAGAA 820

QY 229 GlyIleLeuLeuGlnAlaLeuAlaArgGlnGluGlnGluLysLeuGlnArgValArgLys 248
    |||...|||...|||...|||...|||...|||...|||...|||...|||...|||...|||
Db 821 AGTAGAGTGTCTGACAGCGGATCAGCAGGAGAGAGAAAAAGCTTAACTTCAATCGAGGCC 880

QY 249 LeuIleArgGlnTrpGlyAspHisIleLeuGlnGlySerSerLysLeuValGluSerAlaIle 268
    |||...|||...|||...|||...|||...|||...|||...|||...|||...|||...|||
Db 881 CTCATCCAGCAGTACCAAGAGAGCTGAGCAAGTGCACAAAGCTGTGAAACTGCGCATC 940

QY 269 GlnSerMetGlnGluProGlnMetAlaLeuTrpIleGlnGlnAlaLysGlnLeuIleAsn 288
    |||...|||...|||...|||...|||...|||...|||...|||...|||...|||...|||
Db 941 CAGTCCCTGGAGGAGCCCTGGGAGGCCACCTTCTTGTGACTGACCAAGCACTCATCAAA 1000
  
```

```

QY 289 LysValGlyAlaMetSerLysValGluLeuAlaGlyArgProGluProGlyTrpGlnSer 308
    |||...|||...|||...|||...|||...|||...|||...|||...|||...|||...|||
Db 1001 AGCATTTGGGAAGCTTCCAAAGGCTGCCAGCTGGGAGAACACAGAGGCTTTGAGAAC 1060

QY 309 MetGlnGlnPheSerValSerValGluHisValAlaGluMetLeuArgThrIleAspPhe 328
    |||...|||...|||...|||...|||...|||...|||...|||...|||...|||...|||
Db 1061 ATGACCTCTTACTTGTGATTTAGAGCAGCATAGCAGACGCCCTGAGAACCATTTGACTTT 1120

QY 329 GlnProGlyAlaAlaGlyAspGluGluAspAspMetAla-----LeuAsp 344
    |||...|||...|||...|||...|||...|||...|||...|||...|||...|||...|||
Db 345 GlyGluGlnGlyAsnAlaGlyLeuGlnGlu 354

Db 1169 GAGGAGAGTCCACAGAGGAGGAGAGAA 1198
  
```

## RESULT 12

ABA83058 standard; DNA; 1764 BP.

ABA83058;

05-FEB-2002 (first entry)

Human transcription factor TRFX-85 coding sequence.

Human; transcription factor; TRFX; cell proliferative disease;

autoimmune disease; inflammation; neurological disease;

developmental disorder; cancer; AIDS; infection; cytostatic; anti-HIV;

neuroprotective; antiinflammatory; gene therapy; ds.

Homo sapiens.

W0200172777-42.

04-OCT-2001.

13-MAR-2001; 2001WO-US08117.

13-MAR-2000; 2000US-0189896.

(INCY-) INCYTE GENOMICS INC.

Hillman JL, Baughn MR, Yue H, Lal P, Lu DW, Patterson C;

Azimzai Y, Bandman O, Tang YT, Mathur P, Shah P, Au-Young J;

Reddy R;

WPI; 2001-570896/64.

P-SDB; ABB50234.

Novel transcription factor polypeptides, used to treat diseases

associated with altered activity and expression of TRFX, and to screen

for agents capable of modulating its activity -

Claim 11; Page 311; 327pp; English.

The present sequence is the coding sequence for a human transcription

factor. The transcription factor and its coding sequence are useful in

the diagnosis, treatment and prevention of diseases associated with

altered expression of the transcription factor e.g. cell proliferative,

autoimmune/inflammatory, neurological and developmental disorders. A

number of specific disorders/diseases are given in the specification,

including: arteriosclerosis, cirrhosis, hepatitis, cancers, AIDS,

allergies, anaemia, asthma, autoimmune thyroiditis, bronchitis, atopic

dermatitis, diabetes mellitus, emphysema, Goodpasture's syndrome, gout,

grave's disease, multiple sclerosis, osteoarthritis, pancreatitis,

psoriasis, rheumatoid arthritis, systemic lupus erythematosus, ulcerative

colitis, uveitis, Alzheimer's disease, Huntington's disease, Parkinson's

disease, stroke, and viral, bacterial, fungal and protozoal infections.

Sequence 1764 BP; 473 A; 399 C; 516 G; 376 T; 0 other;

Alignment Scores:



Prod. No.: 2,92e-98 Length: 1764  
 Score: 1098.50 Matches: 210  
 Percent Similarity: 78.00% Conservative: 63  
 Best Local Similarity: 60.00% Mismatches: 66  
 Query Match: 57.45% Indels: 11  
 DB: 22 Gaps: 4

US-09-908-988b-2 (1-366) x ABA83058 (1-1764)

QY 10 LeuLeuGlyAspAlaHisAsnMetAspAsnLeuGluLysGlnLeuIleCysProIleCys 29  
 DB 141 CTGATCCAGATGGAAATCCCATGGAACCTTGAGAAACAGCTGATCTGCCATCTGC 200  
 QY 30 LeuGluMetPheSerLysProValAlaIleLeuProCysGlnHisAsnLeuCysArgLys 49  
 DB 201 CTGGAGATGTTTACCACCCAGCTGATCTTGCCGTCAGACCAACTGTGCCGGAAG 260  
 QY 50 CysAlaAsnAspValPheGlnAlaSerAsnProLeuTrpGlnSerArgLysSerThrThr 69  
 DB 261 TGTGCCATGACATCTTCCAGGCTCAATCCCTACTGAGACCGCGGGCAGCTCAGTG 320  
 QY 70 ValSerSerGlyGlyArgPheArgCysProSerCysArgHisGluValValLeuAspArg 89  
 DB 321 TCCATGCTGTGAGGCCCTTCCCGCTGCCACCTGCCACAGGATGATCATGATCGT 380  
 QY 90 HisGlyValTyrgLysLeuGlnArgAsnLeuValGluAsnIleLeuAspIleTyrgLys 109  
 DB 381 CACGGAGTGAAGGCTCCAGAGAACCTGTGGTGAAGACATCATCGATCTACAAA 440  
 QY 110 GlnGlu---SerSerArgProLeuHisAlaLysAlaGluGlnHisLeuMetCysGluGlu 128  
 DB 441 CAGAGAGCTCCAGTCCGCGCTG-----CAGAAAGGAGATGCCCATGTGCAAGAG 494  
 QY 129 HisGluAspGluLysIleAsnIleTyrgCysLeuSerCysGluValProThrCysSerLeu 148  
 DB 495 CACGAGATGAGAAATCAACATCTACTGCTCAGCTGAGGCTGCCACCTGCTCATG 554  
 QY 149 CysLysValPheGlyAlaHisLysAspCysGluValAlaProLeuProThrIleTyrgLys 168  
 DB 555 TGCAGAGTGTGGATGCATCACAGGCCCTGCGAGTGGCCCATTCGACAGATGCTTCAG 614  
 QY 169 ArgGlnLysSerGluLeuSerAspGlyIleAlaMetLeuValAlaGluAsnAspArgVal 188  
 DB 615 GCACAAAAGATGAGACTGATTAATCTGATCTCATCTGCTGGTGGGGAATGACCGTGTG 674  
 QY 189 GlnAlaValIleThrGlnMetCysGluValLysGlnThrIleGluAspAsnSerArgArg 208  
 DB 675 CAGACCATCATCTACCTGAGCTGAGGATTCCTCCGAGTGACCAAGGAGACAGTCCACAG 734  
 QY 209 GlnLysGlnLeuLeuAsnGlnArgPheGlnThrLeuCysAlaValLeuGluGluArgLys 228  
 DB 735 GTAAAGGAGAGAGCTGAGCAGAGATTTGACAGCTTGTATGCCATCCCGATGAGAAAGAA 794  
 QY 229 GlyLeuLeuLeuGlnAlaLeuAlaArgGluGlnGluLysLeuGlnArgValArgGly 248  
 DB 795 AGTAGTGTCTGACAGCGGATGACGAGGAGGAGGAGAAAGAAAGCTTACTTCATCAGAGGCC 854  
 QY 249 LeuIleArgGlnTyrgLysPheHisLeuGluGlySerSerLysLeuValGluSerAlaIle 268  
 DB 855 CTCAATCCAGAGTACAGAGGAGCTGAGCAAGTCCACAAAGCTGGTGAAGATGCCATC 914  
 QY 269 GlnSerMetGluGlnProGlnMetAlaLeuTyrgLeuGlnGlnAlaLysGluLeuIleAsn 288  
 DB 915 CAGTCCCTGAGAGACCTGGGAGGAGCAGCTTCTCTGTGACTGCCAAGCAATCATCAAA 974  
 QY 289 LysValAlaGlyAlaMetSerLysValGluLeuAlaGlyArgProGluProGlyTyrgLys 308  
 DB 975 AGCATTTGAGAGCTTCCCAAGGCTCCAGCTGGGAGAAACAGAGAGAGCTTTGAGAAC 1034  
 QY 309 MetGluGlnPheSerValSerValGluHisValAlaGluMetLeuArgThrIleAspPhe 328  
 DB 1035 ATGAGCTTCTTACTTGTGATTAGAGACATAGAGAGCCCTGAGAGCCATGACTTT 1094

QY 329 GlnProGlyAlaAlaGlyAspGluGluAspAspMetAla-----LeuAsp 344  
 DB 1095 -----GGGACATGAGGAGAGAGAAATTCATTGAGAAAGATCAGGAA 1142  
 QY 345 GlyGluGluGlyAsnAlaGlyLeuGluGlu 354  
 DB 1143 GAGGAGAGTCCACAGAGAGGAGAGAGAA 1172

RESULT 13  
 AAS25842  
 ID AAS25842 standard; cDNA: 1781 BP.  
 AC AAS25842;  
 XX 07-NOV-2001 (first entry)  
 DE Human cDNA encoding a novel secreted protein, Seq ID 21.  
 XX Human; immunosuppressive; antiarthritic; ss; antirheumatic;  
 KW cyostatic; cardiant; vasotropic; cerebroprotective; noctropic;  
 KW neuroprotective; antibacterial; virucide; fungicide; ophthalmological;  
 KW vulnery; secreted protein; rheumatoid arthritis;  
 KW hyperproliferative disorder; cardiovascular disorder; cardiac arrest;  
 KW cerebrovascular disorder; cerebral ischemia; angiogenesis;  
 KW nervous system disorder; Alzheimer's disease; infection; ocular disorder;  
 KW corneal infection; wound healing; epithelial cell proliferation;  
 KW skin ageing; food additive; preservative; antiproliferative.  
 OS Homo sapiens.  
 XX WO20015322-A2.  
 XX 02-AUG-2001.  
 PD 17-JAN-2001; 2001WO-US01341.  
 EF 31-JAN-2000; 2000US-0179065.  
 XX 04-FEB-2000; 2000US-0180628.  
 PR 24-FEB-2000; 2000US-0184664.  
 PR 02-MAR-2000; 2000US-0186350.  
 PR 16-MAR-2000; 2000US-0189874.  
 PR 17-MAR-2000; 2000US-0190076.  
 PR 18-APR-2000; 2000US-0198123.  
 PR 19-MAY-2000; 2000US-0205515.  
 PR 07-JUN-2000; 2000US-0209467.  
 PR 28-JUN-2000; 2000US-0214886.  
 PR 30-JUN-2000; 2000US-0215135.  
 PR 07-JUL-2000; 2000US-0216647.  
 PR 07-JUL-2000; 2000US-0216880.  
 PR 11-JUL-2000; 2000US-0217487.  
 PR 11-JUL-2000; 2000US-0217496.  
 PR 14-JUL-2000; 2000US-0218290.  
 PR 26-JUL-2000; 2000US-0220963.  
 PR 26-JUL-2000; 2000US-0220964.  
 PR 14-AUG-2000; 2000US-0224518.  
 PR 14-AUG-2000; 2000US-0224519.  
 PR 14-AUG-2000; 2000US-0224519.  
 PR 14-AUG-2000; 2000US-0225213.  
 PR 14-AUG-2000; 2000US-0225214.  
 PR 14-AUG-2000; 2000US-0225266.  
 PR 14-AUG-2000; 2000US-0225267.  
 PR 14-AUG-2000; 2000US-0225268.  
 PR 14-AUG-2000; 2000US-0225270.  
 PR 14-AUG-2000; 2000US-0225447.  
 PR 14-AUG-2000; 2000US-0225757.  
 PR 14-AUG-2000; 2000US-0225758.  
 PR 14-AUG-2000; 2000US-0225759.  
 PR 14-AUG-2000; 2000US-0225759.  
 PR 18-AUG-2000; 2000US-0226279.  
 PR 22-AUG-2000; 2000US-0226681.  
 PR 22-AUG-2000; 2000US-0226688.  
 PR 22-AUG-2000; 2000US-0227182.  
 PR 23-AUG-2000; 2000US-0227009.  
 PR 30-AUG-2000; 2000US-0228924.  
 PR 01-SEP-2000; 2000US-0229287.



```

PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 12-SEP-2000; 2000US-0232081.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 21-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 25-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 01-NOV-2000; 2000US-0241826.
PR 08-NOV-2000; 2000US-024617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246539.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249307.
PR 17-NOV-2000; 2000US-0249308.
PR 17-NOV-2000; 2000US-0249309.
PR 17-NOV-2000; 2000US-0249310.
PR 17-NOV-2000; 2000US-0249311.
PR 17-NOV-2000; 2000US-0249312.

PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249246.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251858.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.

(HUMA-) HUMAN GENOME SCI INC.
Rosen CA, Barash SC, Ruben SM;
WPI: 2001-488783/53.
P-PSDB; AAU15855.

New nucleic acid molecules encoding 461 human secreted proteins for
diagnosing, preventing, treating or ameliorating medical conditions and
used as food additives or preservatives -
XX
PT
PT
XX
PS
Claim 1; SEQ ID NO 21; 980pp; English.

The invention relates to isolated nucleic acid molecules and their
encoded secreted proteins. The nucleic acids and proteins are used to
prevent, treat or ameliorate a medical condition in e.g. humans, mice,
rabbits, goats, horses, cats, dogs, chickens or sheep. They
are also used in diagnosing a pathological condition or susceptibility
to a pathological condition. Antibodies to the proteins can also
be used in alleviating symptoms associated with the disorders and in
diagnostic immunoassays e.g. radioimmunoassays or enzyme linked
immunosorbant assays (ELISA). Disorders which are diagnosed or treated
include autoimmune diseases e.g. rheumatoid arthritis,
hyperproliferative disorders e.g. neoplasms of the breast or liver,
cardiovascular disorders e.g. cardiac arrest, cerebrovascular disorders
e.g. cerebral ischemia, angiogenesis, nervous system disorders e.g.
Alzheimer's disease, infections caused by bacteria, viruses and fungi
and ocular disorders e.g. corneal infection, and many other
disorders listed in the specification. The polypeptides can also
be used to aid wound healing and epithelial cell proliferation, to
prevent skin aging due to sunburn, to maintain organs before
transplantation, for supporting cell culture of primary tissues, to
regenerate tissues and in chemotaxis. The polypeptides can also be used
as a food additive or preservative to increase or decrease storage
capabilities, fat content, lipid, protein, carbohydrate, vitamins,
minerals, cofactors and other nutritional components. The present
sequence encodes a novel secreted protein of the invention.

Alignment Scores:
Pred. No.: 1.44e-97
Score: 1091.50
Percent Similarity: 77.71%
Best Local Similarity: 59.71%
Query Match: 57.09%
DB: 22
Length: 1781
Matches: 209
Conservative: 63
Mismatch: 67
Indels: 11
Gaps: 4

```

US-09-908-988B-2 (1-366) X AAS25842 (1-1781)

OY	10	LeuLeuGlyAspAlaHisAspMetAspIleuGluGlyGlnLeuIleCysProIleCys	29
DB	137	CTGATCCAGAGGGAATCCCATGGGAACCTGGAAACAGCTGATCTGCCCTATCTGC	196
OY	30	LeuGlnMetPheSerLeysProValValIleLeuProCysGlnHisAsnLeuCysArgLys	49
DB	197	CTGGAGATGTTTACCAAGCCAGCTGGTCATCTTGCCGTGCAGACACAACCTGTCCGGAA	256
OY	50	CysAlaAsnAspValPheGlnAlaSerAsnProLeuTyrPheIleSerArgLysTyrThr	69
DB	257	TGTGCCAATGACATCTTCCAGGCTGTGAAATCCCTACTGGACCGCGGGCACTCACTG	316
OY	70	ValSerSerGlyIleArgPheArgCysProSerCysArgHisGlnValValLeuAspArg	89
DB	317	TCCATCTCTGGAGGCCGTTTCCGCTGCCACCTGCCACAGAGGTGATCATGATCGT	376
OY	90	HisGlyValTyrGlyLeuGlnAlaArgAsnLeuLeuValGluAsnIleIleAspIleTyrLys	109
DB	377	CACGAGATTAGGCGCTCGACAGGAACCTGTGGTGGAGAACATCGCATCTCCATCTCAAA	436
OY	110	GlnGlu---SerSerArgProLeuHisAlaTyrAlaGluGlnHisLeuMetCysGlnGlu	128
DB	437	CAGGATGCTCCAGTGGCCCGCTG-----CAGAAGGCGATCACCCCATGTCCAAAGAG	490
OY	129	HisGluAspGluLysIleAsnIleTyrCysLeuSerCysGluValProThrCysSerLeu	148
DB	491	CACGAAGATGAGAAATAATACATCTACTGCTCAGCTGTGAGGTGCCACCTGCTCATG	550
OY	149	CysLysValPheGlyAlaHisLysAspCysGluValAlaProLeuProThrIleTyrLys	168
DB	551	TCCAAAGTCTTTGGGATCCACAAGGCCCTGTGAGGTGGCCCATTTGACAGAGTGTCTTCCAG	610
OY	169	ArgGlnLysSerGluLeuSerAspGlyIleAlaMetLeuValAlaGluAsnAspArgVal	188
DB	611	GGACAAAAGATCACTGATGATATCTGATCTCATGCTGTGGGGGGAATGACCTGTGT	670
OY	189	GlnAlaValIleThrGlnMetGluGluValCysGlnThrIleGluAsnSerArgArg	208
DB	671	CAGACATCATCATCTGACGTGGAGGATCCCTCGATGATGACCAAGGAACAGTCACAG	730
OY	209	GlnLysGlnLeuLeuAsnGlnArgPheGluThrLeuCysAlaValLeuGluArgLys	228
DB	731	GTTAAAGAGAGAGCTGAGCCAGAGAGTTTGACACGTGTATGCCATCTTGATGAGAGAAA	790
OY	229	GlyGluLeuLeuGlnAlaLeuAlaArgGluGlnGluLysLeuGlnArgValArgGly	248
DB	791	AGTGAATCTGTTCGACGGGATCCACGAGGACAGAGAAAAGCTTACTTCATGAGAGCC	850
OY	249	LeuIleArgGlnTyrGlyAspHisLeuGlnGlySerSerLysLeuValGluSerAlaIle	268
DB	851	CTCATCCAGCATGACAGAGCACTGGACAAAGTCCCAAAAGCGTGTGGAAACCTGCATC	910
OY	269	GlnSerMetGluLysProGlnMetAlaLeuTyrLeuGlnGlnAlaLysGluLeuIleAsn	288
DB	911	CACCTCCCTGGACAGAGCTGGGAGAGCCACCTCTCTTGCATCGCCAAAGCAATCATCAA	970
OY	289	LysValGlyAlaMetSerLysValGluLeuAlaGlyArgProGlyLysTyrGlnLys	308
DB	971	AGCATTTGTGAACCTTCCAGGGCTGCCACCTGGGGAAGACAGCAAGGCTTTGAGAAC	1030
OY	309	MetGlnGlnPheSerValSerValGlnHisValAlaGluMetLeuArgThrIleAspPhe	328
DB	1031	ATGGACTTCTTACTTGGATTGATTAAGACACATACAGACGCCGCGAGAGCCATTTGACTT	1090
OY	329	GlnProGlyAlaAlaGlyAspGluGluAspAspMetala-----LeuAsp	344
DB	1091	-----GGAGACAGATGAGGAGAGGAAAGAAATTCATTGAAAGAAAGATCAGGAA	1138
OY	345	GlyGlnGluGlyAsnAlaGlyLeuGlnGlu	354
DB	1139	GAGAGAGAGTCCACAGAAAGGAGGAAGAA	1168



```

Db      292 TGCTCAATGACATCTCCAGGCTGCAAAATCCCTACTGGAACACCGGGGACGCTCAGTGT 351
OY      70  alserSerGlyIyArPhearGcYsProserCyArGHisGluValValLeuAspArgH 90
        |||||||
Db      352 CCAATGCTGGAGGCGTTTCCGCGTCCGACCTGCCGCCACAGAGATGATGATGATGCTC 411
OY      90  tsGlyValIyTgLy-LeuGlnArGAsn-LeuLeuValGluAsnIleIleAspIleTyrIly 109
        |||||||
Db      412 ACCGAGGTACCCGCTGCACAGAACCTTGTGTGTGGAGAACATCATGACATCTACAA 471
OY      109 sgGlnIserSerArPrProLeuHisAlaLysAlaGluGlnHisLeuMetCysGluGlnH 129
        |||||||
Db      472 ACAGAGATGCTTCACATGCGGCCGCTGCAGAGGCCACAGCAACCC-ATGTGCAAGAGCA 530
OY      129 sgIAspGluLysIleAsnIleTyrCysLeuSerCysGluValProThCysSerLeuCy 149
        |||||||
Db      531 CGAAGATGAGAAATCAACATCTACTGTCTCAGTGTGAGTGCACCCGCTCCATGTG 590
OY      149 slyValPheGlyAlaHisLysAspCysGluValAlaProLeuProThrIleTyrIlySar 169
        |||||||
Db      591 CGAGTGTGTGGATCCACAGAGGCTCGAGAGTGGCCCATTTGACAGATGTCTTCCAGGG 650
OY      169 gGlnLysSerGluLeuSerAspGlyIleAlaMetLeuValAlaGlyAsnAspArgValG 189
        |||||||
Db      651 ACAAAAGACTGAAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 710
OY      189 nAlaValIleThrGlnMetGluGluValAlcysGlnThrIleGluAspAsnSerArgArgG 209
        |||||||
Db      711 GACCATCATCTACTGAGCTGAGGAGATTCCTCCGCTGACCAAGAGAACAGTACACACG 770
OY      209 nLys-GlnLeuLeuAsnIleArGpHeGlnThrLeuCysAlaValIleGluGlnIlyArg 229
        |||||||
Db      771 AAATAGAGAGAGCTGAGCCAGAGATTGACACCTGTGATGCTGATGCTGATGATGATGATG 830
OY      229 LysGluLeuGlnAlaLeuValArGpGlnGlnGlnLysLeuGlnArGValArgIlyL 249
        |||||||
Db      831 GTGAGTGTCTGCTCGGATCTCGCAAGACAGAGAAAGAACTTATCTTCTGACGAGCC 890
OY      249 euIleArGpGlnTyrGlyAspHisLeuGlnGlySerSerIysLeuValGluSerAlaIle 269
        |||||||
Db      891 TCATTCACACGATACACGAGGAGCTGAGACAGTCCACAGAGCTGTGGTGAACACTGCCAT 950
OY      269 InsMetGluGlnProGlnMetAlaLeuTyrLeuGlnGlnAlaLysGluLeuIleAsn 289
        |||||||
Db      951 AGTCCCTGAGAGAGCTGGGGAGCCACTTCTCTGACTGCCAACACACTCATCAAAA 1010
OY      289 ySValGlyAlaMetSerIysValGluLeuAlaGlyArGpProGlnIlyTyrGlnSer 309
        |||||||
Db      1011 GCATTTGTGGAAGCTTCCAAAGGCTGCCAGCTGGGGAAGACAGAGCAGGCTTTGAGAACA 1070
OY      309 etGluGlnPheSerValSerValGluHisValAlaGluMetLeuArGpThrIleAspHe 328
        |||||||
Db      1071 TGGACTTCTTTACTTTGATTGAGCATATGACACAGAGCCCTGAGAGCCATTGACTTTG 1130
OY      329 -----GlnProGlyA 332
        |||||||
Db      1131 GGACGATNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNGTCCA 1190
OY      332 laAlaGlyAspGluGln-AspAspAspMetAlaLeuAsp----- 344
        |||||||
Db      1191 CAGAAAGGAGAGAAAGAACACCCAGCTAAAGAGAGCTGATGAATGAGAGGCCCCAGATGC 1250
OY      345 -----GlyGluGlnGlyAsnAlaGlyLeu 352
        |||||||
Db      1251 AGAGAGACTGGAGAGGAGGTGGGAGGGGCCACAGCGGCTT 1289

```

```

XX      Human ORFX ORF840 polynucleotide sequence SEQ ID NO:1679.
DE
XX      Human; open reading frame; ORFX; detection; cytostatic; hepatotropic;
XX      vulnery; antipariatic; antiparkinsonian; neurotropic; neuroprotective;
XX      anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiant;
XX      immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
XX      hypotensive; dermatologic; immunosuppressive; antiinflammatory;
XX      antiviral; antibacterial; antifungal; antirheumatic; antihypertensive;
XX      antianemic; gene therapy; cancer; proliferative disorder; hypertension;
XX      neurodegenerative disorder; osteoarthritis; graft vs host disease;
XX      cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
XX      cholesterol ester storage; systemic lupus erythematosus; infection;
XX      severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
XX      allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
XX      bone damage; cartilage damage; antinflammatory disease; coagulation;
XX      thrombosis; contraceptive; ss.
OS
XX      Homo sapiens.
XX
XX      WO200058473-A2.
XX
XX      05-OCT-2000.
XX
XX      31-MAR-2000; 2000WO-US08621.
XX
XX      31-MAR-1999; 99US-0127607.
XX      02-APR-1999; 99US-0127636.
XX      05-APR-1999; 99US-0127728.
XX      30-MAR-2000; 2000US-0540763.
XX
XX      (CURA-) CURAGEN CORP.
XX
XX      Shinketsu RA, Leach M;
XX
XX      WPI: 2000-602362/57.
XX      P-PSDB: AAB41076.
XX
XX      Novel nucleic acids and peptides derived from open reading frame X,
XX      useful for treating e.g. cancers, proliferative disorders,
XX      neurodegenerative disorders and cardiovascular disease -
XX
XX      Claim 5; Page 1334; 5507pp; English.
XX
XX      AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,
XX      which represent the human ORFX open reading frames 1 to 3161. The ORFX
XX      sequences have activities such as: cytostatic; hepatotropic; vulnery;
XX      antipariatic; antiparkinsonian; neurotropic; neuroprotective;
XX      osteopathic; anticonvulsant; antiarthritic; immunosuppressant;
XX      immunostimulant; cardiant; thrombolytic; coagulant; vasotropic;
XX      antidiabetic; hypotensive; dermatologic; immunosuppressive;
XX      antiinflammatory; antibacterial; antiviral; antifungal; antirheumatic;
XX      antihypertensive; antianemic. The sequences can be used for determining
XX      the presence of or predisposition to, or preventing or treating
XX      pathological conditions associated with an ORFX-associated disorder. The
XX      nucleic acids can be used to express ORFX proteins in gene therapy
XX      vectors. The proteins and nucleic acids may be used to treat cancers,
XX      proliferative disorders, neurodegenerative disorders, osteoarthritis,
XX      graft vs host disease, cardiovascular disease, diabetes mellitus,
XX      hypertension, hypothyroidism, cholesterol ester storage, systemic lupus
XX      erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,
XX      bacterial or fungal infection, malaria, autoimmune disorders, asthma,
XX      allergies, aplastic anaemia, burns, wounds, bone and cartilage damage,
XX      nocturnal haemoglobinuria, antiinflammatory disease; to enhance
XX      coagulation; to inhibit thrombosis; and as a contraceptive.
XX
XX      Sequence 531 BP; 129 A; 143 C; 150 G; 108 T; 1 other;

```

```

Alignment Scores:
Pred. No.: 9.1e-56 Length: 531
Score: 659.50 Matches: 120
Percent Similarity: 88.16% Conservative: 14
Best Local Similarity: 78.95% Mismatches: 15

```

Query Match: 34,498 Indels: 3  
 DB: 21 Gaps: 2  
 US-09-908-988b-2 (1-366) x AAC75285 (1-531)  
 QY 10 LeuLeuGlyAspAlaHisAsnMetAspAsnLeuGluLysGlnLeuIleCysProIleCys 29  
 DB 82 CTGATCCAGATGGGAAATCCCATGGAGAACTTGGAGAAAGAGCTGATCTGCCATCTGC 141  
 QY 30 LeuGluMetPheSerLysProValIleLeuProCysGlnHisAsnLeuCysArgLys 49  
 DB 142 CTGGAGATGTTTACCAAGCAGTGGTCATCTTCCCTGCCAGCAACCTGTGCCGAG 201  
 QY 50 CysAlaAsnAspValPheGlnAlaSerAsnProLeuTrpInsErArgLysSerThrThr 69  
 DB 202 TGTGCCAATGACATCTTCCAGGCTGCAAAATCCCTACTGACACGCGGGCAGCTCAGTG 261  
 QY 70 ValSerSerGlyArgPheArgCysProSerCysArgHisGluValIleLeuAspArg 89  
 DB 262 TCCATGCTCGAGGCGGTTTCCCTGCTACCTGCCGACGAGGATCATGATCATGCTGT 321  
 QY 90 HisGlyValTrpGlyLeuGlnArgAsnLeuValGluAsnIleIleAspIleTrpLys 109  
 DB 322 CACGAGTGTACGGCCTCCAGAGAACTCTGCTGGAGAACATCATCATCATCTACAA 381  
 QY 110 GlnGlu---SerSerArgProLeuHisAlaLysAlaGluGlnHisLeuMetCysGluGlu 128  
 DB 382 CAGGAGTGTCTCCAGTGGCGGCGTG-----CAGAAAGGAGTCAACCCATGTACAGAG 435  
 QY 129 HisGlnAspGluLysIleAsnIleTrpCysLeuSerCysGluValProThrCysSerLeu 148  
 DB 436 CACGAGATGACGAAATCAACATCTACTGCTCAGTGTGAGTGCCACCTGCTCATG 495  
 QY 149 CysLysValPheGlyAlaHisLysAspCysGluVal 160  
 DB 496 TGCAGCTGTTGGATCCACAAGGCTGCGAGGTG 531  
 RESULT 17  
 ABN78104 ID ABN78104 standard; cDNA; 531 BP.  
 XX  
 AC ABN78104:  
 XX  
 DT 08-JUL-2002 (first entry)  
 XX  
 DE Human Interleukin receptor-like ORF3051 cDNA, SEQ ID NO:6101.  
 XX  
 KM Human; ORF: open reading frame; ORFX: drug screening; diagnosis;  
 KM disease monitoring; cytokine; cell proliferation; cell differentiation;  
 KM immune modulation; haematopoiesis regulation; tissue growth;  
 KM angiogenesis; activin; inhibin; chemotactic; chemokine; haemostatic;  
 KM thrombolytic; tumour inhibition; bodily characteristic; fertility;  
 KM behaviour; cancer; proliferative disorder; neurological disorder;  
 KM cardiovascular disease; immune system disorder; organ transplantation;  
 KM tissue growth disorder; tissue regeneration disorder; diabetes mellitus;  
 KM hypothyroidism; cholesterol ester storage disease; infection; vulnery;  
 KM vasotrophic; antipsoriatic; antidiabetic; cytosolic; nucleolar;  
 KM neuroprotective; antithrombotic; antileukemic; antileukemic;  
 KM cardiant; hypotensive; antithyroid; antileukemic; immunomodulator;  
 KM dermatological; analgesic; virucide; antibacterial; fungicide; gene; ss.

PI Leach MD, Shinkets RA;  
 XX WPI: 2002-106200/14.  
 DR P-PSDB; ABP34078.  
 DR  
 XX  
 PT Novel human polypeptides and polynucleotides useful for diagnosing,  
 PT preventing and treating cardiovascular disease, neurodegenerative,  
 PT hyperproliferative disorders and disorders related to organ  
 PT transplantation  
 PT  
 XX  
 PS Claim 1; Page 1775; 2508pp; English.  
 XX  
 CC Sequences ABP1028-ABP35561 represent 4534 novel human proteins  
 CC designated ORF (open reading frame) 1-4534, and sequences ABN75054-  
 CC ABN79587 represent cDNAs encoding them. The invention also encompasses  
 CC polypeptides at least 80% identical to the ORF1-ORF4534 (collectively  
 CC referred to as ORFX) proteins, polynucleotides at least 85% identical to  
 CC the ORFX nucleic acid sequences, vectors and host cells comprising ORFX  
 CC polynucleotides, the recombinant production of ORFX proteins, antibodies  
 CC specific for ORFX proteins, methods of detecting ORFX polynucleotides and  
 CC polypeptides, methods of screening for modulators of ORFX expression or  
 CC activity, and methods of screening individuals for a predisposition to an  
 CC ORFX-associated disorder. The ORFX proteins of the invention have a wide  
 CC range of biological activities, such as cytokine, cell proliferation,  
 CC cell differentiation, immune modulation, haematopoiesis regulation,  
 CC tissue growth, angiogenesis, activin or inhibin activity, chemotactic/  
 CC chemokinetic activity, haemostatic activity, thrombolytic activity,  
 CC receptor/ligand, antiinflammatory activity, tumour inhibition activity,  
 CC and antifibrotic activity, and may also be involved in the determination  
 CC of bodily characteristics, fertility and behaviour. ORFX proteins,  
 CC nucleic acids and antibodies may be used in the treatment of cancers,  
 CC other proliferative disorders such as psoriasis and benign tumours,  
 CC neurological disorders such as epilepsy and Alzheimer's disease,  
 CC cardiovascular diseases, immune system disorders, disorders related to  
 CC organ transplantation, disorders of tissue growth and regeneration,  
 CC diseases such as diabetes mellitus, hypothyroidism, and cholesterol ester  
 CC storage disease, and infectious diseases caused by viral, bacterial,  
 CC fungal and other pathogens. ORFX nucleic acids may also be used as a  
 CC source of primers and probes, in the detection of ORFX genomic sequences  
 CC or transcripts, in the identification and cloning of homologous  
 CC sequences, in genetic diagnosis, and in forensic biology. The ORFX  
 CC nucleic acids may additionally be used to produce transgenic animals  
 CC which may be useful for studying the function and/or activity of ORFX  
 CC protein, and in drug screening. The ORFX proteins may also be used as  
 CC immunogens to generate specific antibodies, which are useful in the  
 CC diagnosis, treatment and monitoring of ORFX-associated diseases.  
 XX  
 SQ Sequence 531 BP; 129 A; 143 C; 150 G; 108 T; 1 other;  
 Alignment Scores:  
 Pred. No.: 9,1e-56 Length: 531  
 Score: 659.50 Matches: 120  
 Percent Similarity: 88.16% Conservative: 14  
 Best Local Similarity: 78.95% Mismatches: 15  
 Query Match: 34,498 Indels: 3  
 DB: 24 Gaps: 2  
 US-09-908-988b-2 (1-366) x ABN78104 (1-531)  
 QY 10 LeuLeuGlyAspAlaHisAsnMetAspAsnLeuGluLysGlnLeuIleCysProIleCys 29  
 DB 82 CTGATCCAGATGGGAAATCCCATGGAGAACTTGGAGAAAGAGCTGATCTGCCATCTGC 141  
 QY 30 LeuGluMetPheSerLysProValIleLeuProCysGlnHisAsnLeuCysArgLys 49  
 DB 142 CTGGAGATGTTTACCAAGCAGTGGTCATCTTCCCTGCCAGCAACCTGTGCCGAG 201  
 QY 50 CysAlaAsnAspValPheGlnAlaSerAsnProLeuTrpInsErArgLysSerThrThr 69  
 DB 202 TGTGCCAATGACATCTTCCAGGCTGCAAAATCCCTACTGACACGCGGGCAGCTCAGTG 261  
 QY 70 ValSerSerGlyArgPheArgCysProSerCysArgHisGluValIleLeuAspArg 89

Dh 262 TCCATGTCGTGAGAGCCGTTCCGTCGCTACCTGCGCCGACGAGGATGATGATGCT 321  
Oy 90 HtsglyValTyrGlyLeuGlnArgAsnLeuValAluAsnIleIleAspIleTyrLys 109  
Dh 322 CACGAGTGTACGGCTGAGAGAACCTGCTGTGAGAACATCATCATCATCAAA 381  
Oy 110 Gtnglu---SerSerArgProLeuHtAlaLysAlaGluGlnHtSlrLeuHtCysGluGlu 128  
Dh 382 CAGGAGTGTCTCCAGTGTGCGCCGCTG-----CAGAGGCGAGTCAACCCCATCTACAGAG 435  
Oy 129 HtsgluAspGluLysIleAsnIleTyrCysLeuSerCysGluValProThrCysSerLeu 148  
Dh 436 CACGAAAGATGAGAAATCAACATCTACTGTCTCAGTGTGAGTGTGCGCCACCTGCTCATG 495  
Oy 149 CysLysValPheGlyAlaHtLysAspCysGluVal 160  
Dh 496 TGCAGAGTGTGTGGATCCACAGAGGCTGCGAGGTG 531  
RESULT 18  
AAS26314  
ID AAS26314 standard; cDNA; 587 BP.  
AC AAS26314;  
XX 07-NOV-2001 (first entry)  
XX  
DE Human cDNA encoding a novel secreted protein, Seq ID 493.  
XX  
KW Human; immunosuppressive; antiarthritic; ss; antirheumatic;  
KW cytosolic; cardiant; vasotropic; cerebroprotective; noctropic;  
KW neuroprotective; antibacterial; virucide; fungicide; ophthalmological;  
KW vulnary; secreted protein; rheumatoid arthritis;  
KW hyperproliferative disorder; cerebral ischaemia; cardiac arrest;  
KW cerebrovascular disorder; cerebral ischaemia; angioneu-  
KW nervous system disorder; Alzheimer's disease; infection; ocular disorder;  
KW corneal infection; wound healing; epithelial cell proliferation;  
KW skin ageing; food additive; preservative; antiproliferative.  
XX  
OS Homo sapiens.  
XX  
PN W0200155322-A2.  
XX  
PD 02-AUG-2001.  
XX  
PF 17-JAN-2001; 2001WO-US01341.  
XX  
XX 31-JAN-2000; 2000US-0179065.  
PR 04-FEB-2000; 2000US-0180628.  
PR 24-FEB-2000; 2000US-0184664.  
PR 02-MAR-2000; 2000US-0186350.  
PR 16-MAR-2000; 2000US-0189874.  
PR 17-MAR-2000; 2000US-0190076.  
PR 18-APR-2000; 2000US-0198123.  
PR 19-MAY-2000; 2000US-0205515.  
PR 07-JUN-2000; 2000US-0209467.  
PR 28-JUN-2000; 2000US-0214886.  
PR 30-JUN-2000; 2000US-0215135.  
PR 07-JUL-2000; 2000US-0216647.  
PR 07-JUL-2000; 2000US-0216880.  
PR 11-JUL-2000; 2000US-0217487.  
PR 11-JUL-2000; 2000US-0217496.  
PR 14-JUL-2000; 2000US-0218290.  
PR 26-JUL-2000; 2000US-0220963.  
PR 26-JUL-2000; 2000US-0220964.  
PR 14-AUG-2000; 2000US-0224518.  
PR 14-AUG-2000; 2000US-0224519.  
PR 14-AUG-2000; 2000US-0225213.  
PR 14-AUG-2000; 2000US-0225266.  
PR 14-AUG-2000; 2000US-0225267.  
PR 14-AUG-2000; 2000US-0225268.  
PR 14-AUG-2000; 2000US-0225270.  
PR 14-AUG-2000; 2000US-0225447.  
XX

PR 14-AUG-2000; 2000US-0225757.  
PR 14-AUG-2000; 2000US-0225758.  
PR 14-AUG-2000; 2000US-0225759.  
PR 18-AUG-2000; 2000US-0226279.  
PR 22-AUG-2000; 2000US-0226681.  
PR 22-AUG-2000; 2000US-0226868.  
PR 22-AUG-2000; 2000US-0227182.  
PR 30-AUG-2000; 2000US-0227009.  
PR 30-AUG-2000; 2000US-0228924.  
PR 01-SEP-2000; 2000US-0229287.  
PR 01-SEP-2000; 2000US-0229343.  
PR 01-SEP-2000; 2000US-0229344.  
PR 01-SEP-2000; 2000US-0229345.  
PR 05-SEP-2000; 2000US-0229309.  
PR 05-SEP-2000; 2000US-0229513.  
PR 06-SEP-2000; 2000US-0230437.  
PR 06-SEP-2000; 2000US-0230438.  
PR 08-SEP-2000; 2000US-0231242.  
PR 08-SEP-2000; 2000US-0231243.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231413.  
PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0232080.  
PR 08-SEP-2000; 2000US-0232081.  
PR 12-SEP-2000; 2000US-0231968.  
PR 14-SEP-2000; 2000US-0232397.  
PR 14-SEP-2000; 2000US-0232398.  
PR 14-SEP-2000; 2000US-0232399.  
PR 14-SEP-2000; 2000US-0232400.  
PR 14-SEP-2000; 2000US-0232401.  
PR 14-SEP-2000; 2000US-0233063.  
PR 14-SEP-2000; 2000US-0233064.  
PR 14-SEP-2000; 2000US-0233065.  
PR 21-SEP-2000; 2000US-0234223.  
PR 21-SEP-2000; 2000US-0234274.  
PR 25-SEP-2000; 2000US-0234597.  
PR 25-SEP-2000; 2000US-0234998.  
PR 26-SEP-2000; 2000US-0235484.  
PR 27-SEP-2000; 2000US-0235834.  
PR 27-SEP-2000; 2000US-0235836.  
PR 29-SEP-2000; 2000US-0236327.  
PR 29-SEP-2000; 2000US-0236367.  
PR 29-SEP-2000; 2000US-0236368.  
PR 29-SEP-2000; 2000US-0236369.  
PR 29-SEP-2000; 2000US-0236370.  
PR 02-OCT-2000; 2000US-0236802.  
PR 02-OCT-2000; 2000US-0237037.  
PR 02-OCT-2000; 2000US-0237038.  
PR 02-OCT-2000; 2000US-0237039.  
PR 02-OCT-2000; 2000US-0237040.  
PR 13-OCT-2000; 2000US-0239395.  
PR 13-OCT-2000; 2000US-0239397.  
PR 20-OCT-2000; 2000US-0240960.  
PR 20-OCT-2000; 2000US-0241221.  
PR 20-OCT-2000; 2000US-0241785.  
PR 20-OCT-2000; 2000US-0241786.  
PR 20-OCT-2000; 2000US-0241787.  
PR 20-OCT-2000; 2000US-0241808.  
PR 20-OCT-2000; 2000US-0241809.  
PR 20-OCT-2000; 2000US-0241826.  
PR 01-NOV-2000; 2000US-0244617.  
PR 08-NOV-2000; 2000US-0244674.  
PR 08-NOV-2000; 2000US-0244675.  
PR 08-NOV-2000; 2000US-0244676.  
PR 08-NOV-2000; 2000US-0246477.  
PR 08-NOV-2000; 2000US-0246478.  
PR 08-NOV-2000; 2000US-0246523.  
PR 08-NOV-2000; 2000US-0246524.  
PR 08-NOV-2000; 2000US-0246525.  
PR 08-NOV-2000; 2000US-0246526.  
PR 08-NOV-2000; 2000US-0246527.  
PR 08-NOV-2000; 2000US-0246528.  
PR 08-NOV-2000; 2000US-0246532.

PR	08-NOV-2000;	2000US-0246609.
PR	08-NOV-2000;	2000US-0246610.
PR	08-NOV-2000;	2000US-0246611.
PR	08-NOV-2000;	2000US-0246613.
PR	17-NOV-2000;	2000US-0249207.
PR	17-NOV-2000;	2000US-0249208.
PR	17-NOV-2000;	2000US-0249209.
PR	17-NOV-2000;	2000US-0249210.
PR	17-NOV-2000;	2000US-0249211.
PR	17-NOV-2000;	2000US-0249212.
PR	17-NOV-2000;	2000US-0249213.
PR	17-NOV-2000;	2000US-0249214.
PR	17-NOV-2000;	2000US-0249215.
PR	17-NOV-2000;	2000US-0249216.
PR	17-NOV-2000;	2000US-0249217.
PR	17-NOV-2000;	2000US-0249218.
PR	17-NOV-2000;	2000US-0249244.
PR	17-NOV-2000;	2000US-0249245.
PR	17-NOV-2000;	2000US-0249264.
PR	17-NOV-2000;	2000US-0249265.
PR	17-NOV-2000;	2000US-0249297.
PR	17-NOV-2000;	2000US-0249299.
PR	17-NOV-2000;	2000US-0249300.
PR	01-DEC-2000;	2000US-0250160.
PR	01-DEC-2000;	2000US-0250391.
PR	05-DEC-2000;	2000US-0251030.
PR	05-DEC-2000;	2000US-0251988.
PR	05-DEC-2000;	2000US-0256719.
PR	06-DEC-2000;	2000US-0251479.
PR	08-DEC-2000;	2000US-0251856.
PR	08-DEC-2000;	2000US-0251868.
PR	08-DEC-2000;	2000US-0251869.
PR	08-DEC-2000;	2000US-0251989.
PR	08-DEC-2000;	2000US-0251990.
PR	11-DEC-2000;	2000US-0254097.
PR	05-JAN-2001;	2001US-0259678.
XX	(HUMA-) HUMAN GENOME SCI INC.	
PA	Rosen CA, Barash SC, Ruben SM;	
PI	WPI; 2001-488783/53.	
PT	P-PDSB; AAU16327.	
DR		
XX	New nucleic acid molecules encoding 461 human secreted proteins for	
PT	diagnosing, preventing, treating or ameliorating medical conditions and	
PT	used as food additives or preservatives -	
PS	Claim 1; SEQ ID No 493; 980bp; English.	
XX		
XX	The invention relates to isolated nucleic acid molecules and their	
CC	encoded secreted proteins. The nucleic acids and proteins are used to	
CC	prevent, treat or ameliorate a medical condition in e.g. humans, mice,	
CC	rabbits, goats, horses, cats, dogs, chickens or sheep. They	
CC	are also used in diagnosing a pathological condition or susceptibility	
CC	to a pathological condition. Antibodies to the proteins can also	
CC	be used in alleviating symptoms associated with the disorders and in	
CC	diagnostic immunoassays e.g. radioimmunoassays or enzyme linked	
CC	immunosorbant assays (ELISA). Disorders which are diagnosed or treated	
CC	include autoimmune diseases e.g. rheumatoid arthritis,	
CC	hyperproliferative disorders e.g. neoplasms of the breast or liver,	
CC	cardiovascular disorders e.g. cardiac arrest, cerebrovascular disorders	
CC	e.g. cerebral ischemia, angiogenesis, nervous system disorders e.g.	
CC	Alzheimer's disease, infections caused by bacteria, viruses and fungi	
CC	and ocular disorders e.g. corneal infection, and many other	
CC	disorders listed in the specification. The polypeptides can also	
CC	be used to aid wound healing and epithelial cell proliferation, to	
CC	prevent skin aging due to sunburn, to maintain organs before	
CC	transplantation, for supporting cell culture of primary tissues, to	
CC	regenerate tissues and in chemotaxis. The polypeptides can also be used	
CC	as a food additive or preservative to increase or decrease storage	
CC	capabilities, fat content, lipid, protein, carbohydrate, vitamins,	
CC	minerals, cofactors and other nutritional components. The present	
CC		

CC	sequence encodes a novel secreted protein of the invention.
Alignment Scores:	
Pred. NO.:	3,14e-51
Score:	614.00
Percent Similarity:	86.00%
Best Local Similarity:	76.67%
Query Match:	32.11%
DB:	22
US-09-908-988B-2 (1-366) x AAS26314 (1-587)	
OY	4 ThrValGIgIyPhelysProLeuLeuGlyAspAlaHisasMetaspAnLeuGluIysGln 23
Db	145 TCTCTGAAATTCACAAATCTTTTTCACAAAGCAGCAGACCATGGTAACTTAAGAAACAA 204
OY	24 LeuIleCysProIleCysLeuGluMePheSerIysProValIleIleauProCysGln 43
Db	205 CTCATCTGTCCCATCTGCTTAAGATGTTCACGAACCCGTGGGATTCCTCTGTCAAG 264
OY	44 HisAsnLeuCysArgIysCysSalaAsnAspValPheGlnIleSeraSnProLeuTrpGln 63
Db	265 CACAACCTGTGTAGGAATGTGGCCAGATGATTTTCCAGGCGCTCTAACCCGGTATTGCC 324
OY	64 SerArgIysSerThrThValSerSerIysGlyArgPheArgCysProSerCysArgHis 83
Db	325 ACAAGAGGAGGTACACCATGGCATCGAGGGGCGGATTCGCTCCATCCCTGTAGACAT 384
OY	84 GluValIleLeuAspArgHisGlyValTyrTrIyLeuGlnArgAsnLeuLeuValGluAsn 103
Db	385 GAAGGAGTTTGGATAGACATGGGGGTATATGGACTTCAGAGAAACCTGTGGTGAATAAT 444
OY	104 IleIleAspIleTyrIysGlnGluSerSerArgProLeuHisAlaIysAlaGluGlnHis 123
Db	445 ATCATTTGACATCTCTACACAGAGATCCACAGGCCA-----GAAAGAAATCCGACCAG 498
OY	124 LeuMetCysGluGluHisGluAspGluIysIleAsnIleTyrCysLeuSerCysGluVal 143
Db	499 CCCAGTGTGCGAGAAACATGAGAGGAGCGCATCAACATCTGTCTGAACTGCGAATA- 557
OY	144 ProThrCysSerLeuCysIysValPheGly 153
Db	558 CCCACTCTCTCTGTGTGCAAGGTGTTGGT 587
RESULT 19	
ABV21919/c	
ID	ABV21919 standard; cDNA; 4345 BP.
XX	AC ABV21919;
XX	13-SEP-2002 (first entry)
DE	Human prostate expression marker cDNA 21910.
KW	Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker.
KM	pharmacogenomic marker; gene; ss.
OS	Homo sapiens.
XX	WO200160860-A2.
XX	23-AUG-2001.
XX	20-FEB-2001; 2001WO-US05171.
XX	17-FEB-2000; 2000US-183319P.
XX	16-MAR-2000; 2000US-189862P.
XX	25-MAY-2000; 2000US-207454P.
XX	09-JUN-2000; 2000US-211314P.
XX	18-JUL-2000; 2000US-219071P.
XX	13-DEC-2000; 2000US-235281P.
PA	(MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.



XX Schlegel R, Endege WO, Monahan JE;  
 DR WPI; 2001-662795/76.  
 XX Novel isolated nucleic acid molecule associated with cancerous state of  
 PT prostate cells and correlating with presence of prostate cancer, useful  
 PT for detecting presence of prostate cancer, stage of prostate cancer -  
 XX  
 PS Claim 1; Page 3733; 11750pp; English.  
 XX The invention relates to an isolated nucleic acid molecule (I) comprising  
 CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the  
 CC specification or its complement. (I) is useful for:  
 CC (a) assessing whether a patient is afflicted with prostate cancer;  
 CC (b) monitoring the progression of prostate cancer in a patient;  
 CC (c) assessing the efficacy of a test compound to inhibit prostate  
 CC cancer in a patient;  
 CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer  
 CC in a patient;  
 CC (e) selecting a composition for inhibiting prostate cancer in a patient;  
 CC (f) assessing the prostate cell carcinogenic potential of a compound;  
 CC (g) determining whether prostate cancer has metastasized in a patient;  
 CC (h) assessing the aggressiveness or indolence of prostate cancer in a  
 CC patient;  
 CC (i) is also useful as a pharmacodynamic or pharmacogenomic marker.  
 XX  
 SQ Sequence 4345 BP; 1321 A; 874 C; 801 G; 1319 T; 30 other;  
 Alignment Scores:  
 Pred. No: 1.12e-31 Length: 4345  
 Score: 427.50 Matches: 112  
 Percent Similarity: 44.30% Conservative: 63  
 Best Local Similarity: 28.35% Mismatches: 127  
 Query Match: 23 Gaps: 10  
 US-09-908-988b-2 (1-366) x ABV21919 (1-4345)  
 QY 19 AsnleuGlulysGlnleuIleucysProIleucysLeuGluMetPheSerIysProVal 38  
 Db 3882 AATATCGAAAGGAGCTCTATTGCCCCAGCATGCAAGAGAGCTGTTTACCACCCA---TTG 38286  
 QY 39 IleuProCysGlnHisAsnleuCysArgIysCysAla----- 51  
 Db 3825 AATCTCCCTTGCCACATAGATGTCGTATTAATGTAAAGAACTCGTGCAGCTC 3766  
 QY 52 -----AsnAspVal----- 54  
 Db 3765 GATGATTCATTCACGATGCGATCGATCGACACTCCATCAAGCAGTCCGACTTCGG 3706  
 QY 55 -----PheGlnAlaSerAsnProIleuTrpGlnSerArg 65  
 Db 3705 CTCCCTCCCTAGTATGATTAATTCACCGAATTAAACAGACGAGGCTGGAAAGCGCAAT 3646  
 QY 66 Gly-----SerThrValSerSerGlyArgPheArgCysProSerCysArg 82  
 Db 3645 TCATTCAGCCCGAGCAACACTGT-----TTCCCTTGCCCTGGCTGGAG 3601  
 QY 83 HisGluValIleuAspArgHisGlyValItyrGlyLeuGlnArgAsnleuLeuValGlu 102  
 Db 3600 CATGATGTGATCTTGGAGAAACGAGAAATCAATGCTGTTCGAAACTTCATTGGAA 3541  
 QY 103 AsnIleuAspIleuTyrLys----- 109  
 Db 3540 ACTATTTGGAAAGATATCGTAAACAGCTAGGCGACCCACACCATTAATGTGTGACCTT 3481  
 QY 110 -----GlnGluSerSerArg----- 114  
 Db 3480 TGTAAACACACACCTCAGAAATCCAAAGAGCTGATGGATGTAGTCAAGTTACTGC 3421  
 QY 115 -----ProIleuHisAlaIleuValGlnHis----- 123

Db 3420 AATGATGCTTCAAAATTCATCACCTTGCGGTACTATAAAGCTCAACATGATATTT 3361  
 QY 124 -----LeuMetCysGluGlnHisGluAspGluLys 133  
 Db 3360 GGTCCAACTACTAATCTTCAAGCCCAAGATTTTAAATGTCCCAAGCAATGAAACAGAGA 3301  
 QY 134 IleAsnIleItyrCysLeuSerCysGluValProThrCysSerLeuCysLysValPheGly 153  
 Db 3300 ATTAACATGACTGTGATTAATATGATAGAGAGCCAGTTCCATCTGTGTAGTTGGGTGCT 3241  
 QY 154 AlaHisLysAspCysGluValAlaProIleuProThrIleTyrLysArgGlnLysSerGlu 173  
 Db 3240 AATCATGCGCAACCAACCTGTAACACATGATGAGACAGTGCCTCAAAACCTTAAGAAAG 3181  
 QY 174 LeuSerAspGlyIleAlaMetLeuValAlaGlnLysAsnAspArgValGlnAlaValIleThr 193  
 Db 3180 CTTTCAAAGATATTTGATTAACCTTATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 3121  
 QY 194 GlnMetGluGluValCysGlnThrIleGluAspAsnSerArgArgGlnLysGlnLeu 213  
 Db 3120 GAATCTAATCTTTAATGAAGAAGAACAGAGCTATGAGAGAGGCTTAAGAAAGCA 3061  
 QY 214 AsnGlnArgPheGluThrLeuCysAlaValIleuGluGluArgLysGlyGluLeuGln 233  
 Db 3060 ATTACACATTTTGAAGAGCTCTTTGAAGTCTGGAAGAGAGAAATCATCTGTTTGAA 3001  
 QY 234 AlaLeuAlaArgGluGlnGluGluLysLeuGlnArgValArgGlyLeuIleArgGlnTyr 253  
 Db 3000 GCATTTGACTCTCTTGAAGAACTAAGATTAACAAATTCCTCAATGGAAGAGTTC 2941  
 QY 254 GlyAspHisLeuGluGluGluGluSerSerLysLeuValGlnSerMetGluGln 273  
 Db 2940 CAGGAGCTCTTGAAGAAACATGGA---CTTGTGGGATATGCTCAAGAGTGTAAAGAG 2884  
 QY 274 ProGlnMetAlaLeuTyrLeuGlnGlnAlaLysGluLeuIleAsnLysValGlyAlaMet 293  
 Db 2883 ACAGATCAGCTCTTCTTGTGACAGACAAAGCAGCTCCACTCAGATATCAGAAAGCC 2824  
 QY 294 SerLysValGluLeuValArgProGluProGluTyrGluSerMetGluGlnPheSer 313  
 Db 2823 ACAGAACTCTTGAAGAGCTTTAGACT---GCAGCTCAGACTCTTTTGAAGACTATGTT 2767  
 QY 314 ValSerValGluHisValAlaGluMetLeuArgThrIleAspPhe 328  
 Db 2766 GTTAATACCTTAACAAACAGAACTTCTTGAGAAATTAATCTTT 2722  
 RESULT 20  
 ID ABV22250 standard; cDNA; 4345 BP.  
 AC ABV22250;  
 DI 13-SEP-2002 (first entry)  
 DE Human prostate expression marker cDNA 22241.  
 KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;  
 KW pharmacogenomic marker; gene; ss.  
 OS Homo sapiens.  
 PN WO200160860-A2.  
 PD 23-AUG-2001.  
 PF 20-FEB-2001; 2001WO-US05171.  
 PR 17-FEB-2000; 2000US-183319P.  
 PR 16-MAR-2000; 2000US-189862P.  
 PR 25-MAY-2000; 2000US-207454P.  
 PR 09-JUN-2000; 2000US-211314P.  
 PR 18-JUL-2000; 2000US-219007P.  
 PR 13-DEC-2000; 2000US-255281P.



XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

XX Schlegel R, Endege WO, Monahan JE;

XX WPI: 2001-662795/7.6.

XX Novel isolated nucleic acid molecule associated with cancerous state of prostate cells and correlating with presence of prostate cancer, useful for detecting presence of prostate cancer, stage of prostate cancer

XX Claim 1: Page 3836-3837; 11750pp; English.

XX The invention relates to an isolated nucleic acid molecule (I) comprising a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the specification or its complement. (I) is useful for:

- CC (a) assessing whether a patient is afflicted with prostate cancer;
- CC (b) monitoring the progression of prostate cancer in a patient;
- CC (c) assessing the efficacy of a test compound to inhibit prostate cancer in a patient;
- CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer in a patient;
- CC (e) selecting a composition for inhibiting prostate cancer in a patient;
- CC (f) assessing the prostate cell carcinogenic potential of a compound;
- CC (g) determining whether prostate cancer has metastasized in a patient;
- CC (h) assessing the aggressiveness or indolence of prostate cancer in a patient;
- CC (I) is also useful as a pharmacodynamic or pharmacogenomic marker.

XX Sequence 4345 BP: 1321 A; 874 C; 801 G; 1319 T; 30 other;

XX Alignment Scores:

Pred. No.: 1.12e-31 Length: 4345  
Score: 427.50 Matches: 112  
Percent Similarity: 44.30% Conservative: 63  
Best Local Similarity: 28.35% Mismatches: 127  
Query Match: 22.36% Indels: 93  
Gaps: 10

US-09-908-988B-2 (1-366) x ABV22250 (1-4345)

OY 19 AsnLeuGlLysGlnLeuIleuLeuLeuProIleuLeuGluMetPheSerIleProValVal 38  
DB 3882 AATATGCAAGGAGAGCTCTTTGCCCCAGCATGAGAGCTGTTTACCACCA---TTG 3826  
OY 39 IleuProCysGlnHisAsnLeuLeuCysArgLysCysAla----- 51  
DB 3825 ATTCTGCCCTTGGCACAATGATGATCTCATTAATGTGTAAGAAGCTCCCTGCTCTC 3766  
OY 52 -----AsnAspVal----- 54  
DB 3765 GATGATTCATTCACGATGCTGGATCAGACAACTCCAAATCAAGCAGTCTCGACTTCG 3706  
OY 55 -----PheGlnAlaSerAsnProLeuTyrPglSerArg 65  
DB 3705 CTCCTCCCTCCCTAGTATGATTAATTTGACCGAATTTAACAGACGAGCTGGAACCGCAAT 3646  
OY 66 Gly-----SerThrThrValSerSerGlyGlyArgPheArgCysProSerCysArg 82  
DB 3645 TCATTCAGCCCGCAGGACACTGTT-----TTCCCTTCCCTGGCTGTGAG 3601  
OY 83 HisGluValValLeuAspArgHisGlyValTyrGlyLeuGlnArgAsnLeuLeuValGlu 102  
DB 3600 CATGATGTGGATCTTGGAGACGAGGAATCAATGCTGTGTTGAAACTTCACCTTGGAA 3541  
OY 103 AsnIleIleAspIleTyrLys----- 109  
DB 3540 ACTATTGTGGAAAGATATGCTCAAGCAGCTAGGGCAGCCACAGCATTTATGTGACCTT 3481  
OY 110 -----GlnGluSerSerArg----- 114  
DB 3480 TGTAAACCAACACACCTCAAGAAATTCACAAAAGCTGCATGCTGTAGTCAAGTTACTGC 3421

OY 115 -----ProLeuHisAlaLysAlaGluGlnHis----- 123  
DB 3420 AATGATGCTTCAAAATTCATCACCCTTGGGGTACATATAAAGCTCAACATGATATGT 3361  
OY 124 -----LeuMetCysGluGlnHisGluAspGluLys 133  
DB 3360 GGTCCAACTACTAATTCAGACCCAAAGATTTTATATGTGCCCCAGACATGAAACAGAGA 3301  
OY 134 IleAsnIleTyrCysLeuSerCysGluValProThrCysSerLeuLysValPheGly 153  
DB 3300 ATTAACATGACTGTCATTAATTAATGAGAGCCGTTTCCATCTGTGTATGGGTGT 3241  
OY 154 AlaHisLysAspCysGluValAlaProLeuProThrIleTyrLysArgGlnLysSerGlu 173  
DB 3240 AATCATGCCAACCCCGCTTAACCATATGAGCAGCGCCACAAAACCTTAAGGAAAG 3181  
OY 174 LeuSerAspGlyIleAlaMetLeuValAlaGlyAsnAspArgValGlnAlaValIleThr 193  
DB 3180 CTTTGAAGGATATGATTAATTAATGAGGAAAGCCAGCTGAAGATCAATATCT 3121  
OY 194 GlnMetGluGluValCysGlnThrIleGluAspAsnSerArgGlnLysGlnLeu 213  
DB 3120 GAACTAAACTTGTATTAAGAAACAGAGTGAATGAGAGAGGCTAAAGAAAGCA 3061  
OY 214 AsnGlnArgPheGluThrLeuCysAlaValLeuGluGluArgGlyGluLeuLeuGln 233  
DB 3060 ATTACACATTTTGAAGAAGCTCTTGAAGTTCTGGAAGAGGAAATCATCTGTTTGA 3001  
OY 234 AlaLeuAlaArgGluGlnGluGluLysLeuGlnArgValArgGlyLeuIleArgGlnTyr 253  
DB 3000 GCAATGTGACTCCTTAAGAAACTAAGATTAGACAAATTCAGACTCAAAATGGAAGGTAC 2941  
OY 254 GlyAspHisLeuGluGluLysSerLeuValGluSerAlaIleGlnSerMetGluGlu 273  
DB 2940 CAGGACTCTTGAAGACATGA---CTTGTGGATATGCTTAAGAAAGCTTAAGAG 2884  
OY 274 ProGlnMetAlaLeuTyrLeuGlnAlaLysGluLeuIleAsnLysValGlyAlaMet 293  
DB 2883 ACAGATCAAGTCTTGTGTCGACAGACGAAAGCAGCTCAGCAATACAGAAAGCC 2824  
OY 294 SerLysValGluLeuAlaGlyArgProGluProGlyTyrGluSerMetGluGlnPheSer 313  
DB 2823 ACAGAAATCTTGAAGAGCTTTAGACCT--GCAGCTCAGACTCTTTTGAAGACTATGT 2767  
OY 314 ValSerValGluHisValAlaGluMetLeuArgThrIleAspPhe 328  
DB 2766 GTTAATACCTCTAAACAAACAGAACTTCTTGGAGAAATTAATCTTT 2722

Search completed: December 3, 2002, 14:10:55  
Job time : 299 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: December 3, 2002, 14:05:55 ; Search time 2707 Seconds

(without alignments)  
3934.846 Million cell updates/sec

Title: US-09-908-988b-2

Sequence: 1 MNFTVGFKPLGDHNNMNL.....EGNAGLEERLDEVEGSLH 366

Scoring table:

BL0SUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2054640 segs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:  
-O=/cgn2\_1/USPFO\_spool/US09080888/runat\_26112002\_112154\_11224/app\_query.fasta\_1.519  
-MODEL=frame+ p2n.model -DEV=xlh  
-DB=GenBdl -QPM=fastap -SUFFIX=ige -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blsum62 -TRANS=human4.0.cdi -List=45  
-DOCALLGN=200 -THR\_SCORE=pcpt -THR\_MAX=100 -THR\_MIN=0 -ALIGN=20 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09080888 @CGN 1.1.1616 @runat.26112002.112154.11224 -NCPU=6 -ICPU=3  
-NO\_XLPRXY -NO\_MMAR -LARGQUERRY -NEG\_SCORES=0 -WATT -LONGLOG -DEV\_TIMOUT=120  
-WARN\_TIMOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl:\*  
1: gb\_ba:\*  
2: gb\_htg:\*  
3: gb\_in:\*  
4: gb\_om:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pr:\*  
10: gb\_ro:\*  
11: gb\_sts:\*  
12: gb\_sy:\*  
13: gb\_un:\*  
14: gb\_vi:\*  
15: em\_da:\*  
16: em\_fun:\*  
17: em\_hum:\*  
18: em\_in:\*  
19: em\_mu:\*  
20: em\_om:\*  
21: em\_or:\*  
22: em\_ov:\*  
23: em\_pat:\*  
24: em\_pl:\*  
25: em\_ph:\*  
26: em\_ro:\*  
27: em\_sts:\*  
28: em\_un:\*

29: em\_vi:\*  
30: em\_htg\_hum:\*  
31: em\_htg\_iny:\*  
32: em\_htg\_other:\*  
33: em\_htg\_mus:\*  
34: em\_htg\_pln:\*  
35: em\_htg\_rtd:\*  
36: em\_htg\_mam:\*  
37: em\_htg\_vrt:\*  
38: em\_sy:\*  
39: em\_htgo\_hum:\*  
40: em\_htgo\_mus:\*  
41: em\_htgo\_other:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	1912	100.0	1431	6	AX418848	AX418848 Sequence
2	1912	100.0	1448	10	AF294790	AF294790 Mus muscu
3	1743	91.2	1329	9	HS291714	AJ291714 Homo sapi
4	1261.5	66.0	1500	6	AX060632	AX060632 Sequence
5	1162	60.7	1750	6	AX418850	AX418850 Sequence
6	1161	60.7	1750	9	BC007750	BC007750 Homo sapi
7	1161	60.7	1925	9	AK091728	AK091728 Homo sapi
8	1161	60.7	2202	9	HS291712	AJ291712 Homo sapi
9	1161	60.7	2634	9	AK091310	AK091310 Homo sapi
10	1157	60.5	1810	9	HS243488	AJ243488 Homo sapi
11	1157	60.5	2098	9	HS243489	AJ243489 Homo sapi
12	1119.5	58.6	1211	9	BC015717	BC015717 Homo sapi
13	1110	58.1	1597	6	AX418852	AX418852 Sequence
14	1109	58.0	1861	10	AY059627	AY059627 Rattus no
15	1102.5	57.7	2097	9	HS291713	AJ291713 Homo sapi
16	1098.5	57.5	1756	9	HS276484	AJ276484 Human MRN
17	1098.5	57.5	1764	6	AX274927	AX274927 Sequence
18	1098.5	57.5	1764	9	AK056942	AK056942 Homo sapi
19	1098.5	57.5	1791	9	AF353673	AF353673 Homo sapi
20	866	45.3	1746	9	AB047601	AB047601 Macaca fa
21	659.5	34.5	531	6	AX313116	AX313116 Sequence
22	624	32.6	199296	2	AC114619	AC114619 Mus muscu
23	624	32.6	209211	2	AC084883	AC084883 Mus muscu
24	624	32.6	216193	10	AC109608	AC109608 Mus muscu
25	601	31.4	1903	9	AF361946	AF361946 Homo sapi
26	594.5	25.9	118847	9	AC013413	AC013413 Homo sapi
27	494.5	25.9	164772	2	AC095184	AC095184 Rattus no
28	439	23.0	1733	9	AF041206	AF041206 Homo sapi
29	439	23.0	2262	10	AF186461	AF186461 Rattus no
30	437	22.9	3334	9	AF035360	AF035360 Homo sapi
31	437	22.9	3352	9	AF230977	AF230977 Homo sapi
32	437	22.9	3422	9	AF041209	AF041209 Homo sapi
33	437	22.9	3452	9	AF230976	AF230976 Homo sapi
34	437	22.9	3452	9	HSXPRE	Y13667 Homo sapien
35	437	22.9	3575	9	AF269101	AF269101 Homo sapi
36	437	22.9	3628	9	AF041208	AF041208 Homo sapi
37	436	22.8	2670	10	MMY14848	Y14848 Mus muscu
38	432	22.6	2004	5	AF269102	AF269102 Gallus ga
39	430	22.5	2511	10	AF026565	AF026565 Mus muscu
40	428	22.4	2298	5	AF374463	AF374463 Gallus ga
41	427.5	22.4	4074	10	HS272269	AJ272269 Homo sapi
42	427	22.3	3348	10	AF186460	AF186460 Mus spret
43	425.5	22.3	2716	9	AB056399	AB056399 Macaca fa
44	424.5	22.2	1875	9	AK026882	AK026882 Homo sapi
45	421.5	22.0	2679	9	AB063010	AB063010 Macaca fa

RESULT 1

#### ALIGNMENTS

AX418848  
 LOCUS AX418848 1431 bp DNA linear PAT 18-JUN-2002  
 DEFINITION Sequence 1 from Patent WO0206318.  
 ACCESSION AX418848  
 VERSION AX418848.1 GI:21523712  
 KEYWORDS  
 SOURCE house mouse.  
 ORGANISM Mus musculus.  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE  
 1 Olson, E.N. and Spencer, J.A.  
 Methods and compositions for stabilizing microtubules and intermediate filaments in striated muscle cells  
 Patent: WO 0206318-A 1 24-JAN-2002.

JOURNAL  
 BOARD OF REGENTS THE UNIVERSITY OF TEXAS SYSTEM (US)

FEATURES  
 source  
 1. 1431  
 /organism="Mus musculus"  
 /db\_xref="taxon:10090"  
 199. 1299  
 /note="unnamed protein product"  
 /codon\_start=1  
 /protein\_id="CAD35455.1"  
 /db\_xref="GI:21523713"  
 /translation="MNFTVGFKKPLGDHNMNDLKEQLICPLEMEKRPVLLPCOH  
 NLCKKANDVFOASNPIMOSRGSTVSGSRCPREHYVLDHGVGLQNLVYE  
 NITDIYKROSSRPLAKADOHLMCEHEDEKINILSCSEVPKSLCKVGAHKDCV  
 APLEPTIKRKSELSDGIMLVACNDROAVITOMEVCOITIDNSRQQLNORE  
 TLCAVLEKREKLQALAREDEKLQVRNGLRQGDHLEGSLSKLSAIOSEEDPM  
 ALTYQQAQKELINKVAMSKVELAGREPECESEMEQFSVEHVAEMLRITDFOPGAAG  
 DEEDDMALDEEGNAGLEERLDYVEGSLH"

CDS  
 338 a 384 c 473 g 236 t

BASE COUNT  
 ORIGIN

Alignment Scores:  
 Pred. No.: 2,46e-146 Length: 1431  
 Score: 1912.00 Matches: 366  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 Gaps: 0

US-09-908-988b-2 (1-366) x AX418848 (1-1431)

QY 1 Metasphehrhrrvaiglypheyproleuenuglyaspahhisasnmetaspasnleu 20  
 Db 199 ATGAACCTTCACGGTGGTTTCAAGCCGCTGCTAGGGGATGGCACAACATGACAACTTG 258

QY 21 Glulysgluleuilecysproillecysleuenglumetpbeserlysprovalvalilleu 40  
 Db 259 GAGAAAGCAGCATATTGGCCCATCTGCTGAGATGTCTCCAAAGCCCGGATGATTTG 318

QY 41 Procyglnhisasnleucysarglyscysalaasnspvalpheelinalaserasnpro 60  
 Db 319 CCGTCCCAACCAACCTGTGCGCAAGTGTGCCAAGAGCTTCCAGGCGCTTAATCCT 378

QY 61 LeutrpclnsrArqglySerThrrhrrvalserSerglylyarphheargcysproser 80  
 Db 379 CTGGGCAATCCCGGGGCTCCACAACGGGTGTTTCAGGAGACGTTCCGATGCCATCT 438

QY 81 Cysarhhisgluvalvalleuasnsparghsglyvaltyrglyleuenglarnasnleu 100  
 Db 439 TGTAGGCGACGAGTGTCTCGTGCACAGGATGCTGTATGGCTGCACGGGAACCTGCTA 498

QY 101 ValgluasnilleleaspriletyrlysglngluserSerarpproleuhisalaalysala 120  
 Db 499 GTGGAGAACATCATTTGACATCTACAAAGCAGGAGCTCTCCGGCCACTGCACGCCAAGCT 558

QY 121 GlulnlnhisleumetCysgluglunhisgluasnpglulysilleasnilletyrcysleuser 140  
 Db 559 GAACAGCACCTCATGTGTGTGAGGAGCACGAGGACGAGAAATCATCAATCTACTGCTGAGC 618

QY 141 CysgluvalProthrcysSerleucyslysalpheelialahislyaspocysgluval 160  
 Db 619 TGGAGAGTGGCCACACTCTCTCTGCAAGTTTTCGGCCGCCACAAGACTGTACAGTG 678

QY 161 AlaProleuProthrlleTyrlYsArgGlnlysserGluleuSeraspolyllaamet 180  
 Db 679 GCCCTCTGCCACCATTTACAAACGCCAGAAAGTGTAGAGGATGGCATGCCGATG 738

QY 181 LeuValAlaGlysnasnspargvalGlnAlaValillethrrglnmetGlnluValysgln 200  
 Db 739 CTGTGGCGGGCAATGACCTGTGTACAGCACTGATCACCCAGATGAGAGGTGTGCCAG 798

QY 201 ThrllgluasnspnsrArqargGlnlysglnleuasnnglnarppheelunthrrleu 220  
 Db 799 ACCATTGAGACACAACGCCGACAGACACAACTGTTAAACAGAGTTCCAGACCTG 858

QY 221 CysAlaValleuenglucuararglysglyleuLeuenglAlaLeuAlaArgglunglu 240  
 Db 859 TCGCGGTTTTGGAGGAGCGCAAGGCGCAACTGCTTCAGCACTGGCCGGAGCAGAG 918

QY 241 Glulysleuclnaryvalaraglyleuillearglnlyrglyasphlsleuengluser 260  
 Db 919 GAGAAATTCAGCGCGGTGCGGGGCTTCATCCGCAAGACGCACTTGGAGGGCTTC 978

QY 261 SerlyleuValgluserAlaillleGlnSerMetgluInuproglnMetAlaleutyrleu 280  
 Db 979 TCAAACTGCTGAGATCCGCGCATTCACATGAGAGAGCCGCAATGCTCTTACTTC 1038

QY 281 GlnGlnAlaIlysgluleuileasnlysalpvalametSerlysalpvalleuAlaIly 300  
 Db 1039 CAGCAGCAAGAGAGCTGATCAACAAGCTGCGGCAATGTGCAAGGTGAGCTGCACAGA 1098

QY 301 ArgpprogluProglytyrtyrGlnSerMetgluInpbeserValserValGlnhisAla 320  
 Db 1099 CGCGCGAGCCAGGCTATGAGACATGAGCAATCTCTGTGACCGTGGACACCTGGCC 1158

QY 321 GluMetleuArghrrlleasppheelnProglyAlaIlaaglyaspGluInaspasp 340  
 Db 1159 GAAATGTTGGAACCATTCGACTTCCACCGCGCGCCCTGGGATGAAGAGATACGAC 1218

QY 341 MetAlaLeuasnpglygluglunGlyasnAlaIlyleuenglunGluarGluasnValpro 360  
 Db 1219 ATGCGTTTGGATGGGAGGAGGCAATGCGGGCTGTGAGAGGAGCGGCTGACCTGCCA 1278

QY 361 GluGlySerGlyleuunhs 366  
 Db 1279 GAAGCTCAGGCTTGAC 1296

RESULT 2  
 LOCUS AF294790 1448 bp mRNA linear ROD 08-SEP-2000  
 DEFINITION Mus musculus RING-finger protein MURF mRNA, complete cds.  
 ACCESSION AF294790  
 VERSION AF294790.1 GI:9945009  
 KEYWORDS  
 SOURCE Mus musculus.  
 ORGANISM Mus musculus.  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE  
 1 Spencer, J.A., Eliazzer, S., Ilaria, R.L. Jr., Richardson, J.A. and Olson, E.N.  
 Regulation of microtubule dynamics and myogenic differentiation by MURF, a striated muscle RING-finger protein  
 J. Cell Biol. 150 (4), 771-784 (2000)

JOURNAL MEDLINE 20411220  
 PUBMED 10953002  
 REFERENCE 2 (bases 1 to 1448)  
 AUTHORS Spencer, J.A. and Olson, E.N.  
 TITLE Direct Submission  
 JOURNAL Submitted (09-AUG-2000) Molecular Biology, UTSW, 6000 Harry Hines Blvd., Dallas, TX 75390-9148, USA  
 FEATURES Location/Qualifiers

source 1. 1448  
 /organism="Mus musculus"  
 /db\_xref="taxon:10090"  
 /tissue\_type="cardiac muscle"  
 199. 1299  
 /codon\_start=1  
 /product="RING-finger protein MURF"  
 /protein\_id="AA03076.1"  
 /db\_xref="GI:9945010"  
 /translation="MNFVGFKPLLDGDAHNMNDLEKOLICPICLEMSPKPVYILPCOH  
 NLCRKANDVFOASNPIMOSRSTSSGGRFRCSCHEVLDHGYGLDRLNLCVE  
 NIIDIKYQESSRPLHAKAEHLMCEBHEHDEINITYCLSCVEPTCSLCVFGAKCEV  
 APIPTITKROSSLSDEIAMLVAGNDROAVITOMEVCOITEDNSRQKOLLNRF  
 TLCAVEERGELLQALAREQEKIQRVGLIROYGDLBESSKLVESAIOSEEPOM  
 ALYQAKELINKVAMSKVELAGRPPEYSGMEQFSVSEHVAEMLRITIDPQGAAG  
 DEEDDMALDGEENAGLEERLDVPEGSGLH"

BASE COUNT 355 a 384 c 473 g 236 t  
 ORIGIN

Alignment Scores:  
 Pred. No.: 2.5e-146 Length: 1448  
 Score: 1912.00 Matches: 366  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 10 Gaps: 0

US-09-908-988b-2 (1-366) x AF294790 (1-1448)

1 MetaspherinValGlyPheLeuLeuGlyAspAlaHisAsnMetAspAsnLeu 20  
 Db 199 ATGACTTCACGGTGGGTTTCAACCGCTGCTAGGGGATGGCCACAACATGACAACTTG 258  
 21 GluLysGlnLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 40  
 Db 259 GAGAGACAGCATTTGCCCATCTGCTGAGAGATGTTCTCCAAAGCCGTTGATCTTG 318  
 41 ProCysGlnHisAsnLeuCysArgLysCysAlaAsnAspValPheGlnAlaSerAsnPro 60  
 Db 319 CCTGGCCAAACAACCTGTGCGCCAGGTGCCAAAGCATGTTCCAGGCTCTATCT 378  
 61 LeuTrpLeuSerArgGlySerThrThrValSerSerGlyGlyArgPheArgCysProSer 80  
 Db 379 CTGTGGCAATCCCGGGGCTCCACAACGGTGTCTTCAGGAGACGTTTCCGATGCCCATCT 438  
 81 CysArgHisGlnValValLeuAspArgHisGlyValTyrGlyLeuGlnArgAsnLeuLeu 100  
 Db 439 TGTAGGACAGAGGTGTCTCGAGACAGCATGTGTCTATGGCTTCACGGAACCTGCTA 498  
 101 ValGlnAsnLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 120  
 Db 499 GTGGAGAACATCATTTGACATCTACACAGAGAGTCTCCGCGCACTCAGCCCAAGGCT 558  
 121 GlnGlnHisLeuMetCysGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 140  
 Db 559 GAACACACACTCATGTGTGAGAGACAGAGAGAGAGATCAACATCTACTGCTCGAGC 618  
 141 CysGlnValProThrCysSerLeuCysLysValPheGlyAlaHisLysAspCysGlnVal 160  
 Db 619 TCGAGAGTGGCCACCTCTCTCTGCAAGGTTTTCGGCGCCCAAGAGACTGTAGGTG 678  
 161 AlaProLeuProThrIleTyrLysArgGlnLysSerGluLeuSerAspGlyTleAlaMet 180  
 Db 679 GCCCTCTGCGCCACCATTTACAAACGCCAGAGAGTGTAGTACGATGATGCGCATG 738  
 181 LeuValAlaGlnLysAspArgValGlnAlaValIleThrGlnMetGlnValCysGln 200  
 Db 739 CTGTGTGGCGGCAATGACCTGTGTGACAGGATGATCCAGATGAGAGAGTGTGCGAG 798  
 201 ThrIleGlnAspAsnSerArgArgGlnLysGlnLeuLeuAsnGlnArgPheGluThrLeu 220  
 Db 799 ACCATTGAGAGCAACAGCCGACAGACAGAACTGTTAAACGAGAGTGTGAGACCTG 858

QY 221 CysAlaValLeuGlnGluArgLysGlyGluLeuGlnAlaLeuAlaArgGlnGln 240  
 Db 859 TGGCGGCTTTTGGAGAGACCGCAAGGCGAAGCTGTTCAAGCATGGCCCGGAGAGAG 918  
 QY 241 GluLysLeuGlnArgValArgGlyLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 260  
 Db 919 GAGAAATTGACGCCGTCGGGGCTCATTCGCCAGTACGGAGACCACTTGGAGGGCTCC 978  
 QY 261 SerLysLeuValGlnSerAlaIleGlnSerMetGluGluProGlnMetAlaLeuTyrLeu 280  
 Db 979 TCAAAGTGTGTGATCCGCGCATCCAGTCCATGAGAGAGCCGCAAGATGGCTCTTACCTC 1038  
 QY 281 GlnGlnAlaLysGlnLeuLeuLeuLeuLeuValGlyAlaMetSerLysValGluLeuAlaGly 300  
 Db 1039 CAGAGCCAAAGAGCGTGTATCAACAAGGTCCGGGCAATGTGCAAGCTGAGCTGCAGGA 1098  
 QY 301 ArgProGluProGlyTyrGlyLysSerMetGlnGlnPheSerValSerValGlnHisValAla 320  
 Db 1099 CGCGCGAGCCAGCGCTATGAGACATGAGCAATCTCTGTGAGCGGTGAGACAGCTGGCC 1158  
 QY 321 GluMetLeuArgThrIleAspPheGlnProGlyAlaAlaGlyAspGluAspAsp 340  
 Db 1159 GAAATGTTGCAACCATGACTTCCAGCGGGCGCCGCTGGGGATGAAGAGATGACGAC 1218  
 QY 341 MetAlaLeuAspGlyGlnGluGlyAsnAlaGlyLeuGlnGluArgLeuAspValPro 360  
 Db 1219 ATGGCTTGGATGGGAGAGAGAGCAATCGGGGCTGTGAGAGAGAGCGGCTGGACCTGCA 1278  
 QY 361 GluLysSerGlyLeuHis 366  
 Db 1279 GAAGCTCAGGCGCTGCAC 1296

RESULT 3  
 HSA291714  
 LOCUS HSA291714 1329 bp mRNA linear PRI 06-JUL-2001  
 DEFINITION Homo sapiens mRNA for RNF30 gene for ring finger protein 30.  
 ACCESSION AJ291714  
 VERSION AJ291714.1 GI:13160387  
 KEYWORDS MURF-3 gene; muscle specific ring finger protein.  
 SOURCE human.  
 ORGANISM Homo sapiens

REFERENCE  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 1 (bases 1 to 1329)  
 AUTHORS  
 Centner,T., Yano,J., Kimura,E., McElhinny,A.S., Pelin,K.,  
 Witt,C.C., Bang,M.L., Trombitas,K., Granzler,H., Gregorio,C.C.,  
 Sorimachi,H. and Labell,S.  
 Identification of muscle specific ring finger proteins as potential  
 regulators of the titin kinase domain

JOURNAL  
 J. Mol. Biol. 306 (4), 717-726 (2001)  
 MEDLINE  
 21140140  
 PUBMED  
 11243782

REFERENCE  
 2 (bases 1 to 1329)  
 CENTNER,T.  
 Direct Submission  
 Submitted (12-FEB-2001) Centner T., Structure and Bio computing,  
 EMBL-Heidelberg, Meyerhofstrasse 1, Heidelberg 69112, Germany  
 Revised by author 22-FEB-2001

REMARK  
 FEATURES  
 source  
 1. .1329  
 Location/Qualifiers  
 1. .1329  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 1. .3  
 /note="putative alternative translation initiation codon"  
 /evidence=not\_experimental  
 49. .1203  
 /gene="RNF30"  
 49. .1203  
 /gene="RNF30"  
 /note="alternative"  
 /codon\_start=1  
 /product="ring finger protein 30"  
 /protein\_id="CAC32842.1"

misc\_feature  
 1. .3  
 /note="putative alternative translation initiation codon"  
 /evidence=not\_experimental  
 49. .1203  
 /gene="RNF30"  
 49. .1203  
 /gene="RNF30"  
 /note="alternative"  
 /codon\_start=1  
 /product="ring finger protein 30"  
 /protein\_id="CAC32842.1"

gene  
 49. .1203  
 /gene="RNF30"  
 49. .1203  
 /gene="RNF30"  
 /note="alternative"  
 /codon\_start=1  
 /product="ring finger protein 30"  
 /protein\_id="CAC32842.1"

CDS  
 49. .1203  
 /gene="RNF30"  
 49. .1203  
 /gene="RNF30"  
 /note="alternative"  
 /codon\_start=1  
 /product="ring finger protein 30"  
 /protein\_id="CAC32842.1"

```

/db_xref="GI:13160389"
/translation="MDNLEKOLICPLEMEFSKPVYLLPCOHNLCKRCANDYFOASNP
LMSRSTVSSGFRPCSCRHVEVDLRHVGGLQRLNLENIIDIKQESSRPLHS
KAOHLMCEHEERKINICYLSCVEPTCLCKVFGAHDCVEAPLPTIKRQKODLT
LPLRECSNTNTYCSLDLPSSDPPIIASONTKIDSELISLMLVAGNRKQVLT
TOMEYCORTEDNSRROKOLITOREPESICAVLEERKELLOLAREOKLORVRLI
RQYGDHLKASSKIVESAIOSMEBPMALYLOAKELINKVMSKVELAGREPQVES
MEQFIVREHVAEMLRITIDFOPASGEEVEVAPDGEESAGFEERPDGP"
join(49..515,642..1203)
/gene="RNF30"
/note="short isoform"
/codon_start=1
/product="ring finger protein 30"
/protein_id="CAC32841.1"
/db_xref="GI:13160388"
/translation="MDNLEKOLICPLEMEFSKPVYLLPCOHNLCKRCANDYFOASNP
LMSRSTVSSGFRPCSCRHVEVDLRHVGGLQRLNLENIIDIKQESSRPLHS
KAOHLMCEHEERKINICYLSCVEPTCLCKVFGAHDCVEAPLPTIKRQKODLT
GIAMLVAGNRKQVLTOMEYCORTEDNSRROKOLITOREPESICAVLEERKELLO
LAREOKLORVRLIROYGDHLKASSKIVESAIOSMEBPMALYLOAKELINKYGA
MSKVELAGREPQVESMEQFIVREHVAEMLRITIDFOPASGEEVEVAPDGEESAGP
FEERPDGP"
BASE COUNT      310 a      375 c      414 g      230 t
ORIGIN
Alignment Scores:
Pred. No.:      1.2e-132      Length:      1329
Score:          1743.00      Matches:      342
Percent Similarity: 87.13%      Conservative: 10
Best Local Similarity: 84.65%      Mismatches: 8
Query Match:    91.16%      Indels:      44
DB:             Gaps:      2
US-09-908-988b-2 (1-366) x HSA291714 (1-1329)
QY      1 MetaspherthaValGlyPheLysProLeuLeuGlyAspAlaHisAsnMetAspAsnLeu 20
      1 ATGAACTTCACAGTGGGTTTCAAGCCGCTGCTAGAGGATGACACAGCATGACCAACTCG 60
      21 GlnLysGlnLeuLleCysProLleCysLeuGluMetPheSerLysProValLleLeu 40
      61 GAGAAAGCAGCTCATCTGCCATCTGCTGCTGAGATGTTCTCCAAACAGATGCTGATCTTG 120
      41 ProCysGlnHisAsnLeuCysArgLysCysAlaAsnAspValPheGlnAlaSerAspPro 60
      121 CCTGCCACACAACTGCTGCCGCAATGTGCCACAGAGCGCTTCCAGGCTCGAATCTT 180
      61 LeuTrpGlnSerArgLysSerThrValSerSerGlyLysArgPheArgCysProSer 80
      181 CTATGGCAGTCCCGGGGCTCCACCACTGTCTTTCAGAGGCGGCTTCCGCTGCCATCG 240
      81 CysArgHisGlnValValLeuAspArgHisGlyValTyrGlyLeuGlnArgAsnLeu 100
      241 TGCAGGCGTGAAGTGTCTGTCGACAGACAGCGGTCTACGGCTTGCAGCAAACTGCTTA 300
      101 ValGlnAsnLleLleAspLleTyrLysGlnGlnLysSerArgProLeuHisAlaLysAla 120
      301 GTGGAGAACATTATGACATTTACAGAGAGATCATCCAGGCGGCTGCACTCCAAAGCT 360
      121 GlnGlnHisLeuMetCysGlnGlnHisGlnAspGlnLysLleAsnLleTyrCysLeuSer 140
      361 GAGCAGCAGCACTCATGTCGAGAGAGATGAAGAAGATCATATTTACTGCTGAGC 420
      141 CysGlnValProThrCysSerLeuCysLysValPheGlnValHisLysAspCysGlnVal 160
      421 TGTGAGGTGCCCACTGCTCTCTGCAAGGTCTTGGTCCCAAGAGACTGTGAGGTG 480
      161 AlaProLeuProThrLleTyrLysArgGlnLys----- 171
      481 GCCCCAGTCCCACTTACAAAGCCAGAAAGAACAGATCTCACTCTGTGCCCAAG 540
      171 ----- 171

```

```

Db      541 CTGAGTGCAGTGGCACAACAACAACTTAAGTGCAGCCTTGATCTCCGAGCTCAAGTAT 600
QY      172 -----SerGlnLeuSerAspGlyLe 178
      601 CCTCCATCTTAGCCTCGCAGAACTAAGATTATGATAGTGAAGCTCACGAGTGCATC 660
QY      179 AlameLleuValAlaGlyAsnAspArgValGlnAlaValLleThrGlnMetGlnLVal 198
      661 GCGATCTGCTGGCAGCAGCATGACCGCGTCCAGAGCATGATCACACAGATGGAGAGCTG 720
QY      199 CysGlnThrLleGlnAspAsnSerArgArgGlnLysGlnLeuLeuAsnGlnArgPheGln 218
      721 TGCAGACTATGAGAGACATAGCCGAGAGCAGAGACAGAGAGAGAGAGAGAGAGAGAGAG 780
QY      219 ThrLeuCysAlaValLeuGlnGlnLysGlnLysGlnLysGlnLysGlnLysGlnLys 238
      781 AGCCTGTGCGCAGTGTGTGAGAGAGCCAAAGGAGAGTGTGTGAGAGAGAGAGAGAGAG 840
QY      239 GlnGlnGlnLysLeuGlnLysValArgGlyLeuLleArgGlnTyrGlyAspHisLeuGln 258
      841 CAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 900
QY      259 GlySerSerLysLeuValGlnSerAlaLleGlnSerMetGlnLysProGlnMetAlaLeu 278
      901 GCTCTCTTAAGCTGTGTGAGAGTCTGCCATTCATGAGAGAGAGAGAGAGAGAGAGAGAG 960
QY      279 TyrLeuGlnGlnAlaLysGlnLysLleAsnLysValGlyAlaMetSerLysValGlnLeu 298
      961 TATCTCCACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1020
QY      299 AlaGlyArgProGlnLysProGlyTyrGlnSerMetGlnLysProGlnMetAlaLysValGln 318
      1021 GCAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1080
QY      319 ValAlaGlnMetLeuArgThrLleAspPheGlnProGlyAlaAlaGlyAspGlnLys 338
      1081 GTGGCCGAATGCTGTGGAGCAGATTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1137
QY      339 AspAspMetAlaLeuAspGlyGlnGlnLysAlaGlnLysGlnLysGlnLysGlnLys 358
      1138 ---GAGGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1194
QY      359 ValProGlnGly 362
      1195 GGGCTTAAGGC 1206
Db      1195 GGGCTTAAGGC 1206
RESULT 4
LOCUS      AX060632      1500 bp      DNA      linear      PAT 22-JAN-2001
DEFINITION      Sequence 54 from Patent WO0078954.
ACCESSION      AX060632
VERSION      AX060632.1 GI:12406053
KEYWORDS
SOURCE
ORGANISM      human.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE
1 (bases 1 to 1500)
AUTHORS      Lal,P., Yue,H., Tang,Y.T., Baughn,M.R., Azimzal,Y. and Tran,B.
TITLE      Human transcriptional regulator proteins
JOURNAL      Patent: WO 0078954-A 54 28-DEC-2000;
Incyte Genomics, Inc. (US)
FEATURES
source
1..1500
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT      401 a      378 c      439 g      282 t
ORIGIN
Alignment Scores:
Pred. No.:      1.84e-93      Length:      1500
Score:          1261.50      Matches:      251
Percent Similarity: 89.04%      Conservative: 17

```

Best Local Similarity: 83.39% Mismatches: 27  
 Query Match: 65.98% Indels: 6  
 DB: 6 Gaps: 4

US-09-908-988b-2 (1-366) x AX060632 (1-1500)

QY 1 MetaspherhvalGlyPheleProleuLeuGlyAspAlaHisAsnMetaspasLeu 20  
 DB 291 ATGAACTTCACAGTGGGTTTCAGCCGGCTGAGGGGATGACACAGCATGACACACTG 350  
 QY 21 GlutysGlnLeuIleCysProIleCysLeuGluMetPheSerLysProValIleLeu 40  
 DB 351 GAGAAAGCAGCTCATCTGCCCCCATCTGCGGAGATGTTCTCCAAACAGAGTGATCTTG 410  
 QY 41 ProCysGlnHisAsnLeuCysArgLysCysAlaAsnAspValPheGlnAlaSerAsnPro 60  
 DB 411 CCGGCCAACACAACTCTGCGGCAATGTGCCAAGAGCTCTCCAGGCGCTCGAATCTT 470  
 QY 61 LeutrgInserArgLysSerThrValSerSerGlyLysArgPheArgCysProSer 80  
 DB 471 CTATGGCAGTCCCGGGGCTCCACCACTGTCTTCAGAGAGCGCTTCCGCTGCCATCG 530  
 QY 81 CysArgHisGlnValValLeuAspArgHisGlnValTyrGlyLeuGlnArgAsnLeu 100  
 DB 531 TGCAGGCGATGAGGTTCCTCGAGACAGACGGGTCTACGGCTCGAGCGAAACCTGCTA 590  
 QY 101 ValGlnAsnIleIleAspIleTyrLysGlnGlnSerSerArgProLeuHisAlaLysAla 120  
 DB 591 GTGGGACATATTATGACATTATACAAAGAGAGTCTCCAGGCGCTGACCTCCAGGCT 650  
 QY 121 GluGlnHisLeuMetCysGlnGluHisGlnAspGlnLysIleAsnIleTyrCysLeuSer 140  
 DB 651 GAGCAGCAGCTCATCTGCGAGAGCATGAAGAAGATCAATATTACTGCTGAGC 710  
 QY 141 CysGlnValProThrCysSerLeuCysLysValPheGlyAlaHisLysAspCysGlnVal 160  
 DB 711 TGTGAGGTGCCACCTGCTCTCTGCAAGGTCTTGTGGTCCCAAGAGCTGTGAGGTG 770  
 QY 161 AlaProLeuProThrIleTyrLysArgGlnLysSerGlnLeuSerAspGlyIleAlaMet 180  
 DB 771 GCCCAGCTGCCACCATTTACAAAGCCCAAGAGAGTGCATGCGATGCGCATGCCGATG 830  
 QY 181 LeuValAlaGlyAsnAspArgValGlnAlaValIleThrGlnMetGluGlnValCysGln 200  
 DB 831 CTGCTGGCAGGCAATGACCGCTGCAAGCAGATGATCACACAGATGAGAGGTGTGCAG 890  
 QY 201 ThrIleGlnAspAsnSerArgArgGlnLysGlnLeuLeuAsnGlnArgPheGlnIleLeu 220  
 DB 891 ACTATCGAGGACAAATAGCCGAGGAGCAGAAAGAGTGTAAACCAAGAGTTTGAAGGCTG 950  
 QY 221 CysAlaValAlaLeuGlnGluArgLysGlyGlnLeuGlnAlaLeuAlaArgGlnGln-G1 240  
 DB 951 TGGCAGAGTGTGAGAGAGGAGCAGAGGTGAGTCTGCAAGCGCTGGCCCGGGAACAGCA 1010  
 QY 240 uGluLysLeuGlnArg---ValArgGlyLeuIleArg---GlnTyrGlyAspHisLeuG1 258  
 DB 1011 GGACAAAGCTTCAACCGCATCCAGCGGAGCTCAATTCGCGTCAATATGGGGAAGCAACTT 1070  
 QY 258 uGlySerSerLysLeuValGlnSerAlaIle-----GlnSerMetGlnIleuProGlnMe 276  
 DB 1071 GGGAAAGCCCTCCCTGCTCAAAAGCCTGGTTGGAGGGGTACATCCAAATCCCAACTTT 1130  
 QY 276 tAlaLeuTyrLeuGlnGlnAlaLysGlnLeuIleAsnLysValGlyAlaMetSerLysVa 296  
 DB 1131 CCACATTTTGGATATACAGAGC---CCAAACGAATAATGTGTGCGGACACTTACTATAAT 1187  
 QY 296 1 296  
 DB 1188 G 1188  
 RESULT 5  
 AX418850 2590 bp DNA linear PART 18-JUN-2002  
 LOCUS AX418850

DEFINITION Sequence 3 from Patent WO0206318.  
 ACCESSION AX418850  
 VERSION AX418850.1 GI:21523714  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 house mouse.  
 Mus musculus.  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL  
 BOARD OF REGENTS THE UNIVERSITY OF TEXAS SYSTEM (US)  
 LOCATION/Qualifiers  
 FEATURES  
 source  
 1. .2590  
 /organism="Mus musculus"  
 /db\_xref="taxon:10090"  
 80. .1717  
 /note="unnamed protein product"  
 /codon\_start=1  
 /protein\_id="CAD35456.1"  
 /db\_xref="GI:21523715"  
 /translation="MSTSLNKSFSKREQQTDNLEKOLICPICLEMFTKPVYIIPCQ  
 NLCKKASDIFQASNPYLPTRGTTVASGGFRCPSCHEVYLDHRYGAGLQRLNVE  
 NIIDYRQESTRPRKLDQPMCEHEEERINILCLNCEVPSLCIKVGAHKDCQVAP  
 LTHVFOROKSELSGDIADVNDROGVISQLEDTKTIIECCRKQODLCEKEDHL  
 YGILEERKTEMTQAITROEKELEHRTLIRKSDHEENKLVSEKISGIFQDEPMAY  
 FIONAKTLIKITVASKARFOMEKLEQGEINSNFTYNIREKILIREIDFSREBE  
 DAGTIDEGREBDAYVEEAEENVOTIASGSEESLEKAEPRQOLPRELDVAPRDLASS  
 PEPFSMPPADVLVITQGEVVPVIGSQQTQSEISGPSAEATDPLFTFSWKQGRKT  
 SSNPCKHSGSLQIQLPLGIEDSSVQSAVEAATNQAAYVSGKSSSTATSOIGF  
 EAPSPQOSALAGSGGVILISQLATSSPSGLINSINE"  
 BASE COUNT 703 a 627 c 625 g 635 t  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 4,47e-85 Length: 2590  
 Score: 1162.00 Matches: 226  
 Percent Similarity: 76.99% Conservative: 55  
 Best Local Similarity: 61.92% Mismatches: 76  
 Query Match: 60.77% Indels: 8  
 DB: 6 Gaps: 4  
 US-09-908-988b-2 (1-366) x AX418850 (1-2590)  
 QY 1 MetaspherhvalGlyPheleProleuLeuGlyAspAlaHisAsnMetaspasLeu 20  
 DB 80 ATGACACTTCTCTGAATTAACAAGTCTTCTCCAAAGAGCAGACCATGGATTAACCTTG 139  
 QY 21 GlutysGlnLeuIleCysProIleCysLeuGluMetPheSerLysProValIleLeu 40  
 DB 140 GAAGAAGCAATGATGTGTCCTGCGCTAGAGATGTTTACAGAGCGCTTGCGCATCTTC 199  
 QY 41 ProCysGlnHisAsnLeuCysArgLysCysAlaAsnAspValPheGlnAlaSerAsnPro 60  
 DB 200 CTTGCGCAGCAACAACTGTGCGAGAAATGTGCCAGATCATCTTCCAGCGCTCTAACCCG 259  
 QY 61 LeutrgInserArgLysSerThrValSerSerGlyLysArgPheArgCysProSer 80  
 DB 260 TACTTACCACCAAGAGAGGAGCACCCTGCGATCAGAGGGGCGGCTTCCGCTCCCTCC 319  
 QY 81 CysArgHisGlnValValLeuAspArgHisGlnValTyrGlyLeuGlnArgAsnLeu 100  
 DB 320 TGCACACATGAGGTGTGTATACAGACATGGGGTCTATAGCATGCGAGAGAACTGCTC 379  
 QY 101 ValGlnAsnIleIleAspIleTyrLysGlnGlnSerSerArgProLeuHisAlaLysAla 120  
 DB 380 GTGGAACAAATATGATATCTACAAAGCAGGAATCCACAGGCCA---GAAAAAAATTTG 436  
 QY 121 GluGlnHisLeuMetCysGlnGluHisGlnAspGlnLysIleAsnIleTyrCysLeuSer 140  
 DB 437 GACCGAGCCC---ATGTGTGAAGAGCATGAAGAGAGAGCATGATCACTATTGTCTGAGAC 493

OY 141 CysGluValProThrCysSerLeuCysLysValPheGlyAlaHisLysAspCysGluVal 160  
 |||||  
 Db 494 TGCGAATGCCACCTGTTCTCTGTCGCAAGGTTTGGCCCATAGAGACTGCCAGTG 553  
 OY 161 AlaProLeuProThrIleTyrLysArgGlnLysSerGluLeuSerAspGlyIleAlaMet 180  
 |||||  
 Db 554 GCTCCCTGACATCATCTGTCAGAGGAGCAAGTCAGAGCTCATGTATGTGCTGTA 613  
 OY 181 LeuValAlaGlyAsnAspArgValGlnAlaValIleThrGlnMetCysGluValCysGln 200  
 |||||  
 Db 614 CTGTGGGAACGACGATAGTCCAGGCTGATCAGCCACTGAGAGCAGCCTGTAAA 673  
 OY 201 ThrIleGluAspAsnSerArgArgGlnLysGlnLeuLeuAsnGlnArgPheGluThr 220  
 |||||  
 Db 674 ACATTGAGGAGCTCTGCGAAGAACGACAGACCTGTGAGAAATTTGTCACCTA 733  
 OY 221 CysAlaValLeuGlnGluValArgLysGlyLeuLeuAlaLeuAlaArgGlnGlu 240  
 |||||  
 Db 734 TACGGCATCTCTGAGAGAGAGAGACTGAAATGACCCAGCATCTGCAACACAGAG 793  
 OY 241 GluLysLeuGlnArgValArgGlyLeuIleArgGlnTyrGlyAspHisLeuGlySer 260  
 |||||  
 Db 794 GAGAACTGGAACTGTCGCACTCTTATCAGGAGATTTCCGATCCTCGAGAACGTA 853  
 OY 261 SerLysLeuValGlnSerAlaIleGlnSerMetGluGluProGlnMetAlaLeuTyrLeu 280  
 |||||  
 Db 854 TCGACAGTTGGTGGAGTCAGAGATCCAGTTGATGATGAGCCGGAATGGCAGATTTCTG 913  
 OY 281 GlnGlnAlaLysGluLeuIleAsnLysValGlyAlaMetSerLysValGluLeuAlaGly 300  
 |||||  
 Db 914 CAGAAATGCCAAGACCCCTGTCACAAAGATCGTGAAGATCGTTCAGATGAGAG 973  
 OY 301 ArgProGluProGlyTyrGlySerMetGluGlnPheSerValSerValGlnHisValAla 320  
 |||||  
 Db 974 AAACCTAGAACAAAGTTTATGATCATGCGCAACTTCACGTAATCTCAATGAGAGAGAA 1033  
 OY 321 GluMetLeuArgThrIleAspPheGlnProGlyAlaAlaGlyAspGluLysAspAsp 340  
 |||||  
 Db 1034 AAATATATCCGTGAAATTTGACTT-----TCTAGAGAGAGAGAGAGAGAGAT 1084  
 OY 341 MetAlaLeuAspGlyGluGluGlyAsnAlaGlyLeuGluGluGluArgLysValPro 360  
 |||||  
 Db 1085 GCAAGAGAAATAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 1135  
 OY 361 GluGlySerGlyLeu 365  
 |||||  
 Db 1136 GAGGCAAGAAATGTT 1150  
 RESULT 6  
 BC007750 1750 bp mRNA linear PRI 12-JUL-2001  
 LOCUS Homo sapiens, clone MGC:12836 IMAGE:4110783, mRNA, complete cds.  
 DEFINITION BC007750  
 ACCESSION BC007750.1 GI:14043531  
 VERSION 1  
 KEYWORDS MGC.  
 SOURCE Homo sapiens.  
 ORGANISM Homo sapiens.  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 1750)  
 Strausberg, R.  
 Direct Submission  
 Submitted (11-MAY-2001) National Institutes of Health, Mammalian  
 Gene Collection (MGC), Cancer Genomics Office, National Cancer  
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
 USA  
 NIH-MGC Project URL: <http://mgc.ncl.nih.gov>  
 Contact: MGC help desk  
 Email: [cgabs-remail.nih.gov](mailto:cgabs-remail.nih.gov)  
 Tissue Procurement: ARCC  
 CDNA Library Preparation: Rubin Laboratory  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
 REMARK COMMENT

DNA Sequencing by: National Institutes of Health Intramural  
 Sequencing Center (NISC),  
 Gaithersburg, Maryland;  
 Web site: <http://www.nisc.nih.gov/>  
 Contact: [nisc\\_mgc@nhgri.nih.gov](mailto:nisc_mgc@nhgri.nih.gov)  
 Shewchenko, Y., Wetherby, K.D., Beckstrom-Sternberg, S.M.,  
 Benjamin, B., Blakesley, R.W., Bouffard, G.G., Brinkley, C.,  
 Dietrich, N.L., Guan, X., Gupta, J., Ho, S.-L., Karlins, E., Legaspi, R.,  
 Lim, M., Maduro, Q.L., Mastello, C., Mastrian, S.D., McLooney, J.C.,  
 McDowell, J., Pearson, R., Snyder, B., Stantilpop, S., Thomas, P.J.,  
 Tjongson, E.E., Touchman, J.W., Tsugeon, C., Vogt, J.L., Walker, M.A.,  
 Zhang, L.-H., and Green, E.D.  
 Clone distribution: MGC clone distribution information can be found  
 through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>  
 Series: IRAL Plate: 17 Row: d Column: 4  
 This clone was selected for full length sequencing because it  
 passed the following selection criteria: Hexamer frequency ORF  
 analysis, Similarity but not identity to protein.  
 Location/Qualifiers  
 1..1750  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="MGC:12836 IMAGE:4110783"  
 /tissue\_type="Muscle, rhabdomyosarcoma"  
 /clone\_id="DH10B-17"  
 /lab\_host="DH10B-R"  
 /note="Vector: pOTB7"  
 190..1500  
 /codon\_start=1  
 /product="Unknown (protein for MGC:12836)"  
 /protein\_id="AAH07750.1"  
 /db\_xref="GI:14043532"  
 /translation="MDNEKOLICPLEMFKPVVILPCQHNLCRKASDITQASNP  
 ILPTRGTTMASGGRPCPSCHREYVLDHGVYGLQRLNVLNIDLYKESIRPERK  
 SDPNCSEHEERINITYLNCVPTCSLKGVAHNDCCVAPLTHFORKSEISDGI  
 AILGVNDVGVISQLEDCTKIECCCKROKOELEKDFYGLILEERNKEMOVT  
 RTQEKLEHVALIKKSDHLENVSLVSGIOFMEPEMAVFLQNAKTLKKISEAS  
 KAFQWEKLEHGVENNHTLVNLRKKIIRIEDFYREDEDEEEREGEGEKREGVG  
 GEAYEVEVEVQCTEFREDEDEPERKASELSQVLEQAPALPAPSSSEPPALPPADA  
 PVTQIGFPAFPQGGAAAPASGSADSEPARHIFSPWLSLNE"  
 ORIGIN  
 BASE COUNT 556 a 389 g 433 t 372 c

## CDS

Alignment Scores:  
 Pred. No.: 3,286-85 Length: 1750  
 Score: 1161.00 Matches: 221  
 Percent Similarity: 75.48% Conservative: 56  
 Best Local Similarity: 60.22% Mismatches: 74  
 Query Match: 60.72% Indels: 16  
 DB: 9 Gaps: 3  
 US-09-908-988b-2 (1-366) x BC007750 (1-1750)  
 OY 1 MetAspPheThrValGlyPheLysProLeuLeuGlyAspAlaHisAsnMetAspAsnLeu 20  
 |||||  
 Db 142 ATGAGCGCATCTCGAATTAACAATCTTTTCCAAAGAGAGAGAGAGAGAGAGAGATTA 201  
 OY 21 GluLysGlnLeuIleLecysProIleLecysLeuGlnMetPheSerLysProValIleLeu 40  
 |||||  
 Db 202 GAGAGACACATCATCTGTCCTTAGAGATGTCACGAAACCTGTGGGATTCCTC 261  
 OY 41 ProCysGlnHisAsnLeuCysArgLysCysAlaAsnAspValPheGlnAlaSerAspPro 60  
 |||||  
 Db 262 CTTGTGCGACACACCTGTGAGGAATGTCCAGCATATTTTCCAGGCTCTTAACCG 321  
 OY 61 LeuTrpGlnSerArgGlySerThrThrValSerSerGlyGlyArgPheArgCysProSer 80  
 |||||  
 Db 322 TATTTGCCACACAAGAGAGGTACACACATGATGATGAGGGGCGGATTCGCCCATTC 381  
 OY 81 CysArgHisGluValIleLeuAspArgHisGlyValTyrGlyLeuGlnArgAsnLeu 100  
 |||||  
 Db 382 TGTAGACATGAAGTGTGTTGATAGACATGGGTATATGACTTCAGAGAACTGCTG 441







```

Db 433 GTGGAATATCATGATGATCTACAGAGGAGTCCACGAGCA-----GAAAGAAA 486
Oy 121 GUGInHsleuMeCysGluGluHsGluaspGluyslleasnIleTyrcysleuSer 140
Db 487 TCCGACACGAGCCATGTGGAGAGACATGAAGAGAGCCGACATCACTACTCTCGAAC 546
Oy 141 CysGluValProthCysSerleuCylysValPheGluAlaHisLysaspCysGluVal 160
Db 547 TCCGACAGTACCACCTGCTCTCTGTGCAAGAGTGTGGTGGTGCACACAAAGCTCCAGGTG 606
Oy 161 AlaProleuprothIleTyrcysArgGlnLysSerGluLeuSeraspGlyIleAlaMet 180
Db 607 GCTCCCTCAGCATCATGTGTTCAGAGACAGAGTGTAGTGTAGTGTAGTGTAGTGTAGTGT 666
Oy 181 LeuValAlaGlyaspArgValAlaGlnAlaValIleThGlnMetGluGluValCysGln 200
Db 667 CTCTGGGCGACAGCATGTGATCCAGGAGGATGATCCAGGAGGATGATCCAGGAGGATG 726
Oy 201 ThrIleGluaspAsnSerArgArgGlnLysGlnLeuLeuaspGlnArgPheGluThrLeu 220
Db 727 ACTATCGAGGATGTTCGACGAAACAGAAACAGAGCTTTGTGAGAGTTGTACTCTG 786
Oy 221 CysAlaValLeuGluGluArgLysGluLeuLeuGlnAlaLeuAlaArgGluGln 240
Db 787 TATGCAATTTTGGAGAGAGAGAAATGAATGACCCAGTATTCACCCAGACCAAGAG 846
Oy 241 GlnLysLeuGlnArgValAlaArgGlyLeuIleArgGlnTyrcysIleGlnGluGlySer 260
Db 847 GAGAAACCTGGAACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 906
Oy 261 SerLysLeuValGlnSerAlaIleGlnSerMetGlnGlnProGlnMetAlaLeuTyrcLeu 280
Db 907 TCAAGCTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 966
Oy 281 GlnGlnAlaLysGlnLeuIleasnLysValAlaGlyAlaMetSerLysValGlnLeuAlaGly 300
Db 967 CAGATGCCAAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1026
Oy 301 ArgProGluProGlyTyrcLeuSerMetGlnGlnPheSerValSerValGlnHsValAla 320
Db 1027 AAAATAGAACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1086
Oy 321 GlnMetLeuArgThrIleaspPhe-----GlnProGly 331
Db 1087 AAGATTAATCGTGAATTTACTTTTACAGAGAGATGAAGATGAAGAGAGAGAGAGAGAG 1146
Oy 332 AlaAlaGlyaspGlnLysaspAspMetAlaLeuaspGlyGlnGluGlnAlaGly 351
Db 1147 GGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1191
Oy 352 LeuGlnGluGlnArgLeuasp 358
Db 1192 GTAGAAGTGAAGAGAGTGAAG 1212

```

RESULT 9  
LOCUS AK091310 2634 bp mRNA linear PRI 15-JUL-2002  
DEFINITION Homo sapiens cDNA FLJ3991 f1s, clone DFNES2007332, moderately similar to Mus musculus RING-finger protein MURF mRNA.  
ACCESSION AK091310  
VERSION AK091310.1 GI:21749650  
KEYWORDS oligo capping; f1s (full insert sequence).  
SOURCE Homo sapiens normal dermal fibroblasts (Neonatal Skin) (NHDF2564)  
CDNA to mRNA, clone: LIB:DFNES2 clone:DFNES2007332.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE  
1 Tashiro, H., Yamazaki, M., Watanabe, K., Kumagai, A., Itakura, S., Fukuzumi, Y., Fujimori, Y., Komiyama, M., Sugiyama, T., Irie, R., Otsuki, T., Sato, H., Wakamatsu, A., Ishii, S., Yamamoto, J., Isono, Y., Kawai, H., Otsu, K., Nishikawa, T., Kimura, K., Yamashita, H.,

Matsuo, K., Nakamura, Y., Sekine, M., Kikuchi, H., Kanda, K., Nagatsuma, M., Murakawa, K., Kanehori, K., Takahashi-Fujii, A., Oshima, A., Sugiyama, A., Kawakami, B., Suzuki, Y., Sugano, S., Nagahari, K., Masuno, Y., Nagai, K. and Isogai, T.  
NEDO human cDNA sequencing project  
Unpublished  
2 (bases 1 to 2634)  
Isogai, T. and Yamamoto, J.  
Direct Submission  
Submitted (04-JUL-2002) Takao Isogai, FLJ Project(HRI Team); 2-6-7 Kazusa-Kamatarai, Kisarazu, Chiba 292-0812, Japan  
(E-mail: genomics@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)  
NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology (RAB); cDNA library construction: Helix Research Institute (HRI) (supported by Japan Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB, HRI, and Biotechnology Center, National Institute of Technology and Evaluation; clone selection for full insert sequencing: HRI and RAB; annotation: HRI and RAB.

FEATURES  
source  
1..2634  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="DFNES2007332"  
/cell\_type="normal dermal fibroblasts (Neonatal Skin)"  
/NHDF2564"  
/clone\_lib="DFNES2"  
/note="cloning vector: pME18SFL3-primary culture, normal dermal fibroblasts"

BASE COUNT 793 a 580 c 675 t  
ORIGIN

Alignment Scores:  
Pred. No.: 5.5e-85 Length: 2634  
Score: 1161.00 Matches: 221  
Percent Similarity: 75.48% Conservative: 56  
Best Local Similarity: 60.22% Mismatches: 74  
Query Match: 60.72% Indels: 16  
DB: 9 Gaps: 3

US-09-908-988b-2 (1-366) x AK091310 (1-2634)

```

Oy 1 MetAsnPheThrValGlyPheLysProLeuLeuGlyAspAlaHisasnMetAspAsnLeu 20
Db 377 ATGAGCGATCTCTGAATTAACAATCTTTTCCAAAGACAGAGACGATGATTAACCTTA 436
Oy 21 GlnLysGlnLeuIleCysProIleCysLeuGlnMetPheSerLysProValValIleLeu 40
Db 437 GAGAGCAACTCATCTGCCCCATCTGCTTAGAGATGTGTACAGAAACCTGTGGTATCTC 496
Oy 41 ProCysGlnHisasnLeuCysArgLysCysAlaAsnAspValPheGlnAlaSerAsnPro 60
Db 497 CCTTGTGACACACACCTGTGTGGAATGTGCGCAGAGATATTTCCAGGCTCTTAACCCG 556
Oy 61 LeuTrpGlnSerArgGlySerThrThrValSerSerGlyIleArgPheArgCysProSer 80
Db 557 TATTGCCCCACAGAGAGAGTACCCATGCGATCAGGGGGCCGATTCGCTGCCATCC 616
Oy 81 CysArgHisGluValValLeuaspArgHisGlyValTyrcGlyLeuGlnArgAsnLeu 100
Db 617 TGTACACATGAAGTGGTTTGTGATAGACATGGGGGTATGTAGACTTCAGAGAGACCTG 676
Oy 101 ValGluasnIleIleAspIleTyrcysGlnGlnLysSerArgProLeuHisAlaLysAla 120
Db 677 GTTGAATAATATATTGACATCTTACAGAGAGAGTCCACAGGCGCA-----GAAAGAAA 730
Oy 121 GlnGlnHisleuMeCysGluGlnLysGluaspGluLyslleasnIleTyrcysleuSer 140
Db 731 TCCGACACGAGCCATGTGCGAGAGACATGAAGAGAGAGCGCATCACTACTCTCGAAC 790
Oy 141 CysGluValProthCysSerleuCylysValPheGluAlaHisLysaspCysGluVal 160

```

```

Db 791 TCGCAAGTACCCACCTGCTCTGTGTCGAAGGTGTTGGTGACACAAAGACTGCCAGGTG 850
Oy 161 AlaProleupProthrlleTyrlsArgGlnLysSerGlnLeuSeraspGlyIleAlaMet 180
Db 851 GCTCCCTCAGCAGCATGTGTCCAGAGACAGAAATCTGAGCTCATGTGATGGCATGCCATC 910
Oy 181 LeuValAlaGlyAsnAspArgValGlnAlaValIleThrGlnMetGlnGluValCysGln 200
Db 911 CTCGGGGCAGACAGATCGATCGATCCAGGAGTGATGATCAGCCGTGGAAGACACTGCMAA 970
Oy 201 ThrIleGlnAspAsnSerArgArgArgGlnLysGlnLeuLeuasnGlnArgPheGluThrLeu 220
Db 971 ACTATCGAGGAATGTTCGAGAAACAGAAACAGACCTTGTGAAGAGTTGATTTACTCTG 1030
Oy 221 CysAlaValLeuGlnGluValArgLysGlnGluLeuGlnAlaLeuAlaArgGlnGlu 240
Db 1031 TATGGCATTTGGAGAGAGAGAAATGAATGACCAACTCATTTACCCGACCCAGAG 1090
Oy 241 GlnLysLeuGlnArgValArgGlyLeuIleArgGlnIleArgGlnIleLysPheLysLeuGlnLys 260
Db 1091 GAGAAACTGGAACATGCTGCTGCTGATCAAAAAGATATCTGATCATTTGGAGACGTC 1150
Oy 261 SerLysLeuValGlnSerAlaIleGlnSerMetGlnGlnProGlnMetAlaLeuTyrlleu 280
Db 1151 TCAAACTTGTTGAGTCAAGAAATTCAGTTTATGATGATGAGCCAGAAATGCGAGTGTCTG 1210
Oy 281 GlnGlnAlaLysGlnLeuIleAsnLysValGlnAlaMetSerLysValGlnLeuAlaGly 300
Db 1211 CAGATGCGCAAAACCTCGTAAAAAATCTCGGAGACATCAAAAGCATTTTCAGATGGAG 1270
Oy 301 ArgProGlnProGlyTyrlleGlnSerMetGlnGlnPheSerValSerValGlnHisValAla 320
Db 1271 AAAATAGAACATGCTGATGAGAACATGACCACTTCACAGTCAACCTCATAGAGAGAA 1330
Oy 321 GlnMetLeuArgThrIleAspPhe-----GlnProGly 331
Db 1331 AAGATTAATACGTGAATATGACTTTTACAGAGAAATGACATGACATGACAGAGAGAAAGGC 1390
Oy 332 AlaAlaGlyAspGlnGlnAspAspMetAlaLeuAspGlyGlnGlnGlnAlaGly 351
Db 1391 GGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1455
Oy 352 LeuGlnGlnGlnArgLysLeuAsp 358
Db 1436 GTAGAGAGTGAAGAGGTAGAA 1456

RESULT 10
LOCUS HSA243488 1810 bp mRNA linear PRI 29-JUN-2001
DEFINITION Homo sapiens titin zinc-finger anchoring protein, 50kDa isoform.
ACCESSION AJ243488
VERSION AJ243488.1 GI:14588845
KEYWORDS alternative splicing; signal transduction; titin zinc-finger anchoring protein; titian.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1810)
REFERENCE 1
AUTHORS Kelly,R., Neubauer,G. and Gautel,M.
TITLE A novel KING finger protein associated with titin kinase
JOURNAL Unpublished
2 (bases 1 to 1810)
AUTHORS Gautel,M.S.
TITLE Direct Submission
JOURNAL Submitted (29-JUN-1999) Gautel M.S., Physikalische Biochemie, Max-Planck-Institut fuer molekulare Physiologie, Otto-Hahn-Strasse 11, Dortmund, 44227, GERMANY
FEATURES
source
1..1810
/organism="Homo sapiens"
/db_xref="taxon:9606"
/cell_type="cardiomyocyte"

```

```

CDS
227..1585
/feature="cardiac muscle"
/function="putative role in signal transduction"
/size="50kDa isoform"
/codon_start=1
/evidence=experimental
/product="titin zinc-finger anchoring protein"
/db_xref="GI:14588846"
/translation="MSASLNKSFKSKQDMNLEKOLICPLCEMTKRPVILPCOH
NICRKAASIFQASNFILPTRGITMASGGRFCPSRHEVVDLRHGVGLQRLVIG
NIDYIKOSTPREKSDPCHEHEERINICLNCIEVLPSCICKYGAHRCQVAP
LHVHFORKSELSGDIALLVGSNDROVGVISLEDCTKTIECCRKOKOECEFDYL
YGLIEERKNEMQVITRTOEKLHVAHLIJKYSIDHLENSKLVESGLOFDEPMAY
FLONAKTLTKKISEASKAFOMEKLEHGEEMNNHTVNLNREKLTIREIDFYRDEDEE
EEEGGGEKEEKEGGEAGEVFEVEVQVEFEFREDENPEKASSELQVELDAARALP
VSSPEPPPLPRPADAPVTOIGFAPLQDGAAPASGSADSEPARHTFSFWLNSL
NE"
BASE COUNT 557 a 420 c 448 g 385 t
ORIGIN

Alignment Scores:
Pred. No.: 7,23e-85 Length: 1810
Score: 1157.00 Matches: 222
Percent Similarity: 76.60% Conservative: 53
Best Local Similarity: 61.84% Mismatches: 74
Query Match: 60.51% Indels: 10
DB: 9 gaps: 3

US-09-908-988b-2 (1-366) x HSA243488 (1-1810)
Oy 1 MetAspPheThrValGlnPheLysProLeuGlnLysPalaHisAsnMetAspAsnLeu 20
Db 227 ATGAGCGCATCTCTGTAATTCAAATCTTTTCCAAAGACGACGACCATGATACTTA 286
Oy 21 GlnLysGlnLeuLecGlyProLecGlyLeuGlnMetPheSerLysProValIleLeu 40
Db 287 GAGAACCACTCATCTGTCCATCTGCTGATGAGATGTTACAGAAACCTGTGATCTGC 346
Oy 41 ProCysGlnHisAsnLeuCysArgLysCysAlaLysAspValPheGlnAlaSerAsnPro 60
Db 347 CCTTGCACACACCTGTGTAGGAAATGTGCCAGTATATTTTCCAGGCCCTGTAAACCG 406
Oy 61 LeuTrpGlnSerArgLysSerThrThrValSerSerGlyArgPheArgCysProSer 80
Db 407 TATTTGCCCAAGAGAGAGTACCACCATGATGCATCAGGGGCCCATTTCCGTCCTCCATCC 466
Oy 81 CysArgHisGlnValValLeuAspArgHisGlyValTyrlleGlnLysGlnAlaGlnLeu 100
Db 467 TGTAGACATGAAAGTGGTTTGGATAGACATGGGGTATATGATGATCAGAGAACTGCTG 526
Oy 101 ValGlnAsnIleLeuAspIleTyrlleGlnGlnLysSerArgProLeuHisAlaLysAla 120
Db 527 GTGGGAATATCATTTACATCTCAAGACAGAGATGCCACAGGCCA-----GAAAAGAAA 580
Oy 121 GlnGlnHisLeuMetCysGlnGlnHisGlnAspGlnLysIleAsnIleTyrlleCysLeuSer 140
Db 581 TCCGACACCCCATGTGCGAGAGAAATGAAAGAGAGGCGCATCATCTATCTGTGAC 640
Oy 141 CysGlnValProThrCysSerLecCysLysValPheGlnAlaHisLysAspCysGlnVal 160
Db 641 TCCGAAAGTACCCACCTGCTCTGTGCAAGGTGTTGGTGACACAAAGATGCGCAGGTG 700
Oy 161 AlaProleupProthrlleTyrlsArgGlnLysSerGlnLeuSeraspGlyIleAlaMet 180
Db 701 GCTCCCTCAGCAGCATGTGTCCAGAGACAGAAATCTGAGCTCATGTGATGGCATGCCATC 760
Oy 181 LeuValAlaGlyAsnAspArgValGlnAlaValIleThrGlnMetGlnGluValCysGln 200
Db 761 CTCGGGGCAGACAGATCGATCCAGGAGTGATGATGACCACTGGAAGAGACACTGCMAA 820
Oy 201 ThrIleGlnAspAsnSerArgArgGlnLysGlnLeuLeuasnGlnArgPheGluThrLeu 220

```

Db 821 ACTATGAGGAATGTTGACAGAAACAGAAACAGAGCTTTGTGAGAGATTGATTACCTG 880  
Oy 221 CysAlaValLeuGluGluArgLysGlyGluLeuLeuGluAlaLeuAlaArgGluGlu 240  
Db 881 TATGCGATTGTTGGAGGAGGAGAAATGAAATGACCAATCTATTACCCAGACCAAGAG 940  
Oy 241 GluLysLeuGluArgValArgGlyLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 260  
Db 941 GAGAAACCTGGAACATGCTCCGTCTGTATCAAAAGATATTCTGATCTTTGGAGAACGTC 1000  
Oy 261 SerLysLeuValGluSerAlaLeuGlnSerMetGluGluProGlnMetAlaLeuTyrLeu 280  
Db 1001 TCAAGTTGTTGAGTACGAGCAATTCATTATGATGAGAGCAAAATGCGAGTCTTCTG 1060  
Oy 281 GlnGlnAlaLysGluLeuLeuLeuLysValGlyAlaMetSerLysValGluLeuAlaGly 300  
Db 1061 CAGAAATGCCAAACCTGCTGTAATAAAATCTCGAAGACATCAAAAGCATTCAGATGAG 1120  
Oy 301 ArgProGluProGlyTyrLeuSerMetGluGlnPheSerValSerValGluHisValAla 320  
Db 1121 AAAATAGAACATGCGTATGAGAACATGACCACTTCACACTCACTCATTAGAGAGAA 1180  
Oy 321 GluMetLeuArgThrIleAspPheGlnProGlyAlaAlaGlyLysAspGluLysAsp 340  
Db 1181 AAGATAATACGTGAATATGACTTTTAC-----AGAGAAATGAGAGAT 1222  
Oy 341 MetAlaLeuAspGlyGluGluGlyAsnAlaGlyLeuGluGluGluGluGluGluGlu 359  
Db 1223 GAA-----GAAAGAGAGAGAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1273  
RESULT 11  
LOCUS HSA243489 2098 bp mRNA linear PRI 29-JUN-2001  
DEFINITION Homo sapiens titin zinc-finger anchoring protein, 60kda isoform.  
ACCESSION AJ243489.1 GI:14588847  
VERSION AJ243489.1 GI:14588847  
KEYWORDS alternative splicing; signal transduction; titin zinc-finger anchoring protein; titian.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 2098)  
Kellly,R., Neubauer,G. and Gautel,M.  
TITLE A novel RING finger protein associated with titin kinase  
JOURNAL Unpublished  
AUTHORS Gautel,M.S.  
TITLE Direct Submmission  
JOURNAL Submitted (29-JUN-1999) Gautel M.S., Physikalische Biochemie,  
Max-Planck-Institut fuer molekulare Physiologie, Otto-Hahn-Strasse  
11, Dortmund, 44227, GERMANY  
FEATURES  
source location/Qualifiers  
1..2098  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/cell\_type="cardiomyocyte"  
/tissue\_type="cardiac muscle"  
227..1873  
/function="putative role in signal transduction"  
/note="60kda isoform"  
/evidence-experimental  
/product="titin zinc-finger anchoring protein"  
/protein\_id="CAC43020.1"  
/db\_xref="GI:14588848"  
/translation="MSASLNYSFSSKEQOTMDNLEKOLICPICLEMFTKPVLLPCOH  
NLCRCASDIPQASNPYLPTRGSTMASGFRPCSRHVVLDNRHVGSLORNLVIG  
NIDYIKOSTRPEKSDOPMCBEHEERINITYLCEWPTGSCIKFEGAHKCOVAP  
LTHVROKSELSDEGAILVSGNDROQVITISOLEMCKTEIECCRRKQKQCEFDL  
YGLIERKEMTOVITRTOEKLEHVALIKTSIDLENSKLVESSTQIMDEPEMAY  
FLQNAKTLTKRISEASKAFQMEKIEHGYENMHFTVNLNREKIIRIDIFYRDEDEE  
EEEGEGEKEEGEVEGAEVVEEVENVOTEFPEDEENPEKASELSQVELQAAFGALP

VSSEPPALPPADAPVTOGEVPTGSEQTTESETPVPAAEADLPLEPSMYGOT  
RKATNPCTGSEGLGIGPGSEDSVVRKAEVAAASERAAVSQKETSAPATMSQ  
IGFAPPLGGAAAPASGSGADSEPARHIFESFWLSNLE"  
BASE COUNT 632 a 497 c 534 g 435 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 8,72e-85 Length: 2098  
Score: 1157.00 Matches: 222  
Percent Similarity: 76.60% Conservative: 53  
Best Local Similarity: 61.84% Mismatches: 74  
Query Match: 60.51% Indels: 10  
Gaps: 3  
US-09-908-988b-2 (1-366) x HSA243489 (1-2098)  
Oy 1 MetAspPheThrValGlyPheLysProLeuGluGlyAspAlaHisAsnMetAspAsnLeu 20  
Db 227 ATGACGCGATCTCTGAATATACAAATCTTTTCCAAAGACAGACAGACATCGATTA 286  
Oy 21 GluLysGlnLeuLeuLecysProLleCysLeuGluMetPheSerLysProValAlaIleLeu 40  
Db 287 GAGAGCAACTCATCTGTCCTCCATCTGCTTAGAGATGTCAGAAACCTGTGTGATTCTC 346  
Oy 41 ProCysGlnHisAsnLeuLecysArgLysCysAlaAsnAspValPheGlnAlaSerAspPro 60  
Db 347 CCTTCTACGACCAACCTGTGTGAGAAATGTGCCAGGATATTTTCCAGGCTTAAACCC 406  
Oy 61 LeuTrpGlnSerArgLysSerThrThrValSerSerGlyLysArgPheArgCysProSer 80  
Db 407 TATTTGCCCAAGAGAGAGTACCCACCATGCAATCAGGGGGCGATTCGCTCCCATCC 466  
Oy 81 CysArgHisGluValValLeuAspArgHisGlyValTyrGlyLeuGlnArgAsnLeuLeu 100  
Db 467 TGTAGACATGAAAGTGTGTTGATGATACATGGGGTATATGAGATTTCAGAGAACTGTG 526  
Oy 101 ValGluAsnIleLeuAspIleTyrLysGlnLysSerSerArgProLeuHisAlaLysAla 120  
Db 527 GTGGAAATATATCATTTGACACTCTACAAAGCAGAGTCCACCAAGCCA-----GAAAGGAA 580  
Oy 121 GlnGlnHisLeuMetCysGluGluHisGluAspGluLysIleAsnIleTyrCysLeuSer 140  
Db 581 TCCGACCAAGCCCATGTGTCCAGAGAACATGAAAGACGACCATCAACATCTCTGCAAC 640  
Oy 141 CysGluValProThrCysSerLeuLecysLysValPheGlyAlaHisLysAspCysGluVal 160  
Db 641 TCGAAGTACCCACCTGCTCTGTGCAAGGTGTTGGTCACACAAAGACTCCAGGTG 700  
Oy 161 AlaProLeuProThrIleTyrLysArgGlnLysSerGluLeuSerAspGlyIleAlaMet 180  
Db 701 GCTCCCTCATTCTCATCTGTTCAGAGACAGAAAGTCTGAGCTCATGATGCAATCGGCATC 760  
Oy 181 LeuValAlaGlyAsnAspArgValGlnAlaValIleThrGlnMetGluValCysGln 200  
Db 761 CCGTGGCGAGCAACAGATGAGTCCAGGAGTACAGCACCATGGAAGACACCTCGCAA 820  
Oy 201 ThrIleGluAspAsnSerArgArgGlnLysGlnLeuLeuAsnGlnPheGluThrLeu 220  
Db 821 ACTATGAGGAATGTTGCAAGAAACAGAAACAGAGCTTTGTAGAGATTGATTACCTG 880  
Oy 221 CysAlaValLeuGluGluArgLysGlyGluLeuLeuGluAlaLeuAlaArgGluGlu 240  
Db 881 TATGCGATTGTTGGAGGAGGAGAAATGAAATGACCAATCTATTACCCAGACCAAGAG 940  
Oy 241 GluLysLeuGluArgValArgGlyLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 260  
Db 941 GAGAAACCTGGAACATGCTCCGTCTGTATCAAAAGATATTCTGATCTTTGGAGAACGTC 1000  
Oy 261 SerLysLeuValGluSerAlaLeuGlnSerMetGluGluProGlnMetAlaLeuTyrLeu 280  
Db 1001 TCAAGTTGTTGAGTACGAGCAATTCATTATGATGAGAGCAAAATGCGAGTCTTCTG 1060  
Oy 281 GlnGlnAlaLysGluLeuLeuLysValGlyAlaMetSerLysValGluLeuAlaGly 300

Db 1061 CAGATGCGCAAAACCTGCTAAATAAAATCTCGAAGCAACCAAGCATTCAGATGAG 1120  
Oy 301 ArgProGluProGlyTyrGluSerMetGluGlnPheSerValSerValGluHisValAla 320  
Db 1121 AAATAAGAACATGCGTATAGAACATGAACTTCAACCTCAACCTCAATAGAGAA 1180  
Oy 321 GluMetLeuArgThrIleAspPheGlnProGlyAlaIleAlaGlyAspGluAspAsp 340  
Db 1181 AAGATAAATACGTGAATTTACTTTTAC-----AGAGAGATGAGAGAT 1222  
Oy 341 MetAlaLeuAspGlyGluGluGlnAsnAlaGlyLeuGluGluGluArgLeuAspVal 359  
Db 1223 GAA-----GAAGAAGAGAGAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGTG 1273  
RESULT 12  
BC015717  
LOCUS BC015717 1211 bp mRNA linear PRI 11-OCT-2001  
DEFINITION Homo sapiens, similar to ring finger protein 28, clone MGC:17320  
IMAGE:3922363, mRNA, complete cds.  
ACCESSION BC015717  
VERSION BC015717.1 GI:16041696  
KEYWORDS MGC.  
SOURCE Homo sapiens.  
ORGANISM Homo sapiens.  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 1211)  
AUTHORS Strausberg, R.  
TITLE Direct Submission  
JOURNAL Submitted (09-OCT-2001) National Institutes of Health, Mammalian  
Gene Collection (MGC), Cancer Genomics Office, National Cancer  
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
USA  
REMARK NIH-MGC Project URL: http://mgc.nci.nih.gov  
COMMENT Contact: MGC help desk  
Email: cgapbs-remail.nih.gov  
Tissue Procurement: ATCC/DCTD/DTF  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)  
DNA Sequencing by: Genome Sequence Centre,  
BC Cancer Agency, Vancouver, BC, Canada  
Info@gscc.bc.ca  
Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,  
Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,  
Leticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo  
Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven  
Nees, Pawan Pandoh, Anna-Lisa Pridhu, Parvaneh Saeeidi, Jacqueline  
Schelin, Duane Smalhus, Michael Smith, Lorraine Spence, Jeff Stott,  
Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy,  
George Yang, Scott Zuyderduyn, Marco Marra.  
Clone distribution: MGC clone distribution information can be found  
through the I.M.A.G.E. Consortium/LNLN at: http://Image.lnl.gov  
Series: IRAK Plate: 15 Row: 0 Column: 19  
This clone was selected for full length sequencing because it  
passed the following selection criteria: matched mRNA gi: 14211914.  
FEATURES  
source  
1. 1211  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="MGC:17320 IMAGE:3922363"  
/tissue\_type="Skin, melanotic melanoma."  
/clone\_lib="NIH\_MGC\_72"  
/lab\_host="DH10B"  
/note="Vector: PCMV-SPORT6"  
135. 1175  
/codon\_start=1  
/product="Similar to ring finger protein 28"  
/protein\_id="AAH15717.1"  
/db\_xref="GI:16041697"  
/translation="MDYRSSLSIQDNPMENLEKOLICPICLENFTKPVVILPCOHNLIC  
RKCANIDFOANPNPWTSSRSSYSMSGGRFPTCRHEVIMDRHGVYGLGNLIVENTII

DIYKOCSSREPLQKSHPMCKEHDEKINIYCLTCEVPTSCMKVEGIIHKACEVAPLO  
SVQVQGGKTEINNMISMLVAGNDRPOTITITQEDSRRTYRKESHQKEELSOXFTPLIA  
IIDEKSELIQRTTOEDKLSFTLEALIQQOEDLDSKLVETAIQSLDEPGATFL  
LTKOLIKRSIVLEASRGQGLKEQFEGNMDFITLDEHIALDALRAIDGAEDDEDD  
VDFRKQRTBEDD"  
BASE COUNT 336 a 302 c 345 g 228 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 4.84e-82 Length: 1211  
Score: 1119.50 Matches: 211  
Percent Similarity: 81.38% Conservative: 60  
Best Local Similarity: 63.36% Mismatches: 57  
Query Match: 58.55% Indels: 5  
DB: 9 Gaps: 3  
US-09-908-988B-2 (1-366) x BC015717 (1-1211)  
Oy 10 LeuLeuGlyAspAlaHisAsnMetAspAsnLeuGluGlnLeuIleCysProIleCys 29  
Db 153 CTGATCAGGATGGGATGCCATGCCATGAGACTTGGAGAGACAGCTGATCGCCATCTGC 212  
Oy 30 LeuGluMetPheSerLeuProValValIleLeuProCysGlnHisAsnLeuCysArgLys 49  
Db 213 CTGGAGATGTTTACCAAGCCAGGTGTCATCTTGCCTGCCAGCACACCTGTGCCGGAG 272  
Oy 50 CysAlaAsnAspValPheGlnAlaSerAsnProLeuTrpGlnSerArgLysThrThr 69  
Db 273 TGTGCCAATGACATCTTCCAGGCTGCAAAATCCCTACTGGACACGCCGGGAGCTCAGTG 332  
Oy 70 ValSerSerGlyArgPheArgCysProSerCysArgHisGluValValLeuAspArg 89  
Db 333 TCCATGTCTGAGAGCCCTTTCCTGCCCTGCCACCTGCCCCACGAGCTGATGATGCT 392  
Oy 90 HisGlyValTyrGlyLeuGlnArgAsnLeuValGluAsnIleIleAspIleTyrLys 109  
Db 393 CACGGAGTGTACGGCCTCGACAGAGAACCTGCTGTGGAGACATCATCGACATTCACAA 452  
Oy 110 GlnGlu---SerSerArgProLeuHisAlaLysAlaGluGlnHisLeuMetCysGluGlu 128  
Db 453 CAGGAGTGTCCAGTCGCGCGCTG-----CAGAAGGCACTCACCCCATGTGCAAGAG 506  
Oy 129 HisGluAspGluLysIleAsnIleTyrCysLeuSerCysGluValProThCysSerLeu 148  
Db 507 CACGAGATAGAGAAATACACATCTACTGTTCACGTGTGAGAGTCCACCGCTGCATG 566  
Oy 149 CysLysValPheGlyAlaHisLysAspCysGluValAlaProLeuProThrIleTyrLys 168  
Db 567 TGCMAAGTGTGTGGATCCACAAAGGCTGCGAGGTGCCCATTCACAGATGCTTCCAG 626  
Oy 169 ArgGlnLysSerGluLeuSerAspGlyIleAlaMetLeuValAlaGlyAsnAspArgVal 188  
Db 627 GCACAAAGACCTAAGCTGAATACGTGATCTCCATGCTGGGGGGGAAATGACCGTGTG 686  
Oy 189 GlnAlaValIleThrGlnMetGluGluValValCysGlnThrIleGluAspAsnSerArg 208  
Db 687 CAGACATCATCATCTACGCTCGAGAGATTCCTCGTAGAGACACAGAAACAGTATCCAG 746  
Oy 209 GlnLysGlnLeuLeuAsnGlnArgPheGlnThrLeuCysAlaValLeuGluGluArgLys 228  
Db 747 GTAAGGAAGAGCTGAGCCAGAGATTGACACGTTGATGCCATCTGATGAGAGAA 806  
Oy 229 GlyGluLeuLeuGlnAlaLeuAlaArgGluGlnGluGluLysLeuGlnArgValArgLys 248  
Db 807 AGTGAATTCCTGACGCGGATCACGACGACGACGAGAGAAAGCTTACCTCATCAGGCC 866  
Oy 249 LeuIleArgGlnTyrGlyAspHisLeuGlnLysLeuGlySerLeuValGlnSerAlaIle 268  
Db 867 CTCATCCACAGACAGACAGGACGCTGACACAGTCCACAAAGCTGCTGGAAGTCCATC 926  
Oy 269 GlnSerMetCysGluProGlnMetAlaLeuTyrLeuGlnGlnAlaLysGluLeuIleAsn 288  
Db 927 CAGTCCCTGAGACAGCCTGCGGGAGGACCACTTCTCTTGAAGTCCAGCACTATCAAA 986





JOURNAL Science 294 (5547), 1704-1708 (2001)  
 MEDLINE 21578247  
 PUBMED 11679633  
 REFERENCE 2 (bases 1 to 1861)  
 AUTHORS Nunez, L. and Glass, D.J.  
 JOURNAL Direct Submission  
 Submitted (12-OCT-2001) Muscle Research, Regeneron Pharmaceuticals,  
 777 Old Saw Mill River Road, Tarrytown, NY 10591, USA  
 FEATURES  
 SOURCE  
 1. 1861  
 /organism="Rattus norvegicus"  
 /strain="Sprague-Dawley"  
 /db\_xref="taxon:10116"  
 95. 1150  
 /note="Murfl: monomeric E3 ubiquitin ligase"  
 /codon\_start=1  
 /product="muscle ring finger protein 1"  
 /protein\_id="AA16405.1"  
 /db\_xref="GI:16444656"  
 /translation="MDYKSLIPDGNAMENLEKOLICPICLEMFTRPVILPCOHNLC  
 RKANDIRQANPWTNRGSSVMSGSRFPCSRHEVIMDRHGVYIGORNLIVENTII  
 DYOEGSSRPLQKSHMCKEHEDEKINICCTCPVCSICKVGAHACAEVAPQ  
 SIPOGKRELNSCIEMLYVANDRYOTIIQLEDSCHRYTENSQVKELSHKEDALYA  
 ILDKRSRLORITQDEKIDPTEALIIQYREQLKSTKLVBTAIQSLDEPGATPL  
 LSAKPLKSLVASKGQLKTEQGFENNMYFTLNLMEIAELRALIDFGTDEBEFEFE  
 EEEEDQDEGVSTEGHQ"  
 BASE COUNT 490 a 433 c 534 g 404 t  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 5.94e-81 Length: 1861  
 Score: 1109.00 Matches: 216  
 Percent Similarity: 75.07% Conservative: 64  
 Best Local Similarity: 57.91% Mismatches: 79  
 Query Match: 10 Indels: 14  
 Gaps: 5  
 US-09-908-988b-2 (1-366) x AY059627 (1-1861)  
 QY 1 MetanpethrvalIglYpethrvalSerSerGlyYARpPhaYqYsProSer 20  
 Db 95 ATGATTTAAATCGGC-----TTGATTCGCGACGGAATGCTATGGAACACTG 145  
 QY 21 GlulysgluIleuIleCysProIleCysLeuGluMetPheSerIlyProValIleleu 40  
 Db 146 GAGAGAGCGCTCATCTGCCCCATCTGCTTGAGATGTACCAAGCTGGGTCATCTG 205  
 QY 41 ProCysGlnHISanLeuCysArgIlyCysAlaAsnAspValPheGlnAlaSerAsp 60  
 Db 206 CCTGCCAGCAACCTCTGCGGGAAGTGTGCCAACGACATCTTCCAGGCTGCCAATCCC 265  
 QY 61 LeuTrpGlnSerArgGlySerThrThrValSerSerGlyYARpPhaYqYsProSer 80  
 Db 266 TACTGGACCAACCGGGGCGGTGCCATGCTCGAGAGTGTTCCTCCCTGCCCTCG 325  
 QY 81 CysArgHISgluValIleuAspArgHISgluValTyrglyLeuGlnArgAsnLeu 100  
 Db 326 TGGCCCATGTAAGTCATGATGACCGGATGGGGTGTACGGTCTGACAGAGAACTGCTG 385  
 QY 101 ValGlnsnIleIleaspIleTyrglysgIngu---SerSerArgProleuHISAlaLys 119  
 Db 386 GTGGAACATCATCTGACATCTACAGACGGAATCTCCAGTCGCGCCCTG-----CAG 439  
 QY 120 AlaGlnIleHISLeuMetCysGlnIleHISgluAspGluLysIleAsnIleTyrcysLeu 139  
 Db 440 AAAGCAGCCACCGCATGTGACAGAACACGAAACGAAATACATCTACTGCTC 499  
 QY 140 SerCysgluValProthCysSerLeuCysIlyValPheGlyAlaHISLysAspCysGlu 159  
 Db 500 ACGTGCGAGGTGCTACTTCTCTTGTGCAAGGTGTTCGGGGCTCACACGAGCCTGTGAA 559  
 QY 160 ValAlaProleuProthIleTyrglysrArgGlnLysSerGlnLeuSerAspGlyIleAla 179

Db 560 GTTGGCCCCCTTACAACATCTTCCAGAGACAGAACTGAACACTGACATTCCTCC 619  
 QY 180 MetLeuValAlaGlyAsnAspArgValGlnAlaValIleThrGlnMetGluLys 199  
 Db 620 ATGCTGTGGCAGAGGAGACGACCGAGTTCAGACTATCATCTGCTGAGAGACTCTCC 679  
 QY 200 GlnThrIleGlnAspAsnSerArgArgGlnLysGlnLeuLeuAsnGlnArgHISgluThr 219  
 Db 680 CGAGTGCACCAAGAAACAGCCACACGAGGTGAGGAGGAGGACACCAAGTTGACGCC 739  
 QY 220 LeuCysAlaValleuGlnGluArgIlyGlyLeuLeuGlnAlaLeuAlaArgGln 239  
 Db 740 CTCTAGCCCATCTCGACGAGAAAGAGATGCTGCTGCACCGGATCCTACAGAGCAG 799  
 QY 240 GluGlnLysLeuGlnArgValArgGlyLeuIleArgGlnTyrglyAspHISLeuGln 259  
 Db 800 GAGAGAAAGCTTGACCTTCATCTGAGAGCCCTGCTCCAGTACCGAGACAGACGTTGAAAG 859  
 QY 260 SerSerLysLeuValGlnSerAlaIleGlnSerMetGlnGluProGlnMetAlaLeuTy 279  
 Db 860 TCGACCAAGCTTGTGAAACACCATCATCTCCCTGATGAGCCCGAGAGGCCACTTC 919  
 QY 280 LeuGlnGlnAlaLysGlnLeuIleAsnLysValGlyAlaMetSerLysValGlnLeuAla 299  
 Db 920 CTCTGAGTCCCAAGCCGCTCATCAAGACGATTTGAGAACTTCAAGGCGTCCACGCTG 979  
 QY 300 GlyArgProGluProGlyTyrglySerMetGlnGlnPheSerValSerValGlnHisVal 319  
 Db 980 GGGAAAGACAGAACAGAGCTTTGAGAACATGACATCTTACTCTGAATTGAAACACATA 1039  
 QY 320 AlaGlnMetLeuArgThrIleAspPhe-----GlnProGlyAlaAlaLys 335  
 Db 1040 GCAGAGCCCTTGAGGGCCATCTGACTTTGGACACATGACGAGGAGGACTTACTGAGAGG 1099  
 QY 336 GluGlnAspAspAsp-----MetAlaLeuAspGlyGlnGluGlyAsnAlaGly 351  
 Db 1100 GAGAGAGAGAGGATCAAGAAAGAGCGGTGCCACAGAGGAGACCAATGAAGACTGGA 1159  
 QY 352 LeuGlnGluGlnArgLeuAspValProGlnGlyserGly 364  
 Db 1160 TGAGTGAAGAGCCCTCTGATGCAAGAGATGGGGAGGGG 1198  
 RESULT 15  
 LOCUS HSA291713  
 DEFINITION Homo sapiens mRNA for RNF28 gene for ring finger protein 28.  
 ACCESSION AJ291713.1 GI:13171050  
 VERSION AJ291713.1  
 KEYWORDS ring finger protein 28; RNF28 gene.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 2097)  
 Centner, T., Yano, J., Kimura, E., McElhinny, A.S., Pelin, K.,  
 Wilt, C.C., Bang, M.L., Trombitas, K., Granzler, H., Gregorio, C.C.,  
 Sorimachi, H. and Labelle, S.  
 Identification of muscle specific ring finger proteins as potential  
 regulators of the titin kinase domain  
 J. Mol. Biol. 306 (4), 717-726 (2001)  
 JOURNAL MEDLINE 21140140  
 PUBMED 11243782  
 REFERENCE 2 (bases 1 to 2097)  
 Centner, T.  
 TITLE Direct Submission  
 JOURNAL Submitted (12-FEB-2001) Centner T., Structure and Biocomputing,  
 EMBL-Heidelberg, Meyerhofstrasse 1, Heidelberg 69112, Germany  
 REMARK Revisited by author 22-FEB-2001  
 FEATURES  
 SOURCE  
 1. 2097  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 misc\_feature 436..438



```

gene
CDS
    /note="putative alternative translation initiation codon"
    /evidence="not_experimental"
    /gene=".1497"
    /gene="RNF28"
    /gene="RNF28"
    /gene="RNF28"
    /note="short isoform"
    /codon_start=1
    /product="ring finger protein 28"
    /protein_id="CAC33173.1"
    /db_xref="GI:13171051"
    /translation="MENTEROLICPLEMEPTKRVVILPCOHNTCRKANIFQANP
    YTRSRSSVSMSCGRPCPTCRHEVINDRNGVITGLRNLVENTIDIDYKQCSSPLQ
    KSHPEKKEHEDERINYLCTCEVPTCSKCVFQIHAKCEVAPQSVFQKTELNLC
    ISMLVANDRNVQTIITQLEDNRVTKENSQKVELSQKEDTLAIIDERSKELORI
    TOEKEKRSIFEALIOQYOLDKSTKIVETAIIDFGDEGGATFELITAKOLIKSIVA
    SKGCOGKTQGFENMDFITLDEHIDALRAIDFGDEDEEERIEEDQDEESTEG
    KEEGHO"
BASE COUNT      565 a      460 c      604 g      468 t
ORIGIN
Alignment Scores:
Pred. No.:      2,33e-80      Length:      2097
Score:          1102.50      Matches:      211
Percent Similarity: 78.00%      Conservative: 62
Best Local Similarity: 60.29%      Mistakes:      66
Query Match:    57.66%      Indels:      11
DB:             Caps:      4
US-09-908-988b-2 (1-366) x HSA291713 (1-2097)
QY 10 LeuLeuGlyAspAlaHisAsnMetAspAsnLeuGluysGlnLeuIleCysProIleCys 29
    |||::: ||| ::| |||:::|||||:|||||:|||||:|||||:|||||:|||||:
Db 454 CTGATCCAGATGGAGATCCATCGAGAACTTGAGAGACGAGCTGATGCCCTTCATGCG 513
    |||:::|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 30 LeuGluMetPheSerIysProValValIleLeuProCysGlnHisAsnLeuCysArgIlys 49
    |||:::|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 514 CTGGAGATGTTTACCAAGCAGTGGTCAATCTTGCCGCCAGCACACCTGCGCGGAG 573
    |||:::|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 50 CysAlaAsnAspValPheGlnAlaSerAsnProLeuTPGInserArgIlySerThrThr 69
    |||:::|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 574 TGTGCCAATACATCTTCCAGGCTGCAAAATCCCTACTGGACACGCGGAGCGCTGAGTG 633
    |||:::|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 70 ValSerSerIylarPheArgCysProSerCysArgHisGluValValLeuAspArg 89
    |||:::|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 634 TCCATGCTGTGAGGCGCTTCCGCTGCCGACCTGCCGACGAGGTGATGATGATGATGATG 693
    |||:::|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 90 HisGlyValIlyrGlyLeuGlnArgAsnLeuValGluAsnIleIleAspIleIlyrIlys 109
    |||:::|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 694 CAGGAGTGTACGAGCGCTGCAAGAGACCTGCTGGTGGAGAACATCATGACATCATCAAA 753
    |||:::|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 110 GlnGlu---SerSerArgProLeuHisAlaIlyAsnGluGlnHisLeuMetCysGluGlu 128
    |||:::|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 754 CAGAGAGTGTCCAGTGGCGCGCTG-----CAGAAAGGCACTCACCCCATGTGCAAGAG 807
    |||:::|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 129 HisGluAspGluIlyIleAsnIleIyTrCysLeuSerCysGluValProThrCysSerLeu 148
    |||:::|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 808 CACGAGATAGAGAAATCAACATCTACTGTCTCAGTGTGAGTGCCACCTGCTCCATG 867
    |||:::|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 149 CysIysValPheGlyAlaHisIlyAspCysGluValAlaPheProIleThrIleIlyrIlys 168
    |||:::|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 868 TGCAGAGTGTGGAGTCCACAAAGGCTGCGAGGTGGCGCCCATGTCAGAGTGTCTTCAG 927
    |||:::|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 169 ArgGlnIlySerGluLeuSerAspGlyIleAlaMetLeuValAlaGlyAsnAspArgVal 188
    |||:::|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 928 GCACAAAGAGTACGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 987
    |||:::|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 189 GlnAlaValIleThrGlnMetGluGlnValCysGlnThrIleGluAspAsnSerArgArg 208
    |||:::|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 988 CAGACCATCATCATGCTGAGGATTCCTCCGAGTACGACAGAGACAGTACACAG 1047
    |||:::|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 209 GlnIlyGlnLeuLeuAsnGlnArgPheGlnThrLeuCysAlaValIleGluGluArgIlys 228
    |||:::|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

```

Db 1048 GTAAGGAGAGAGCTGAGCCAGAACTTGTAGACGTTGTATGCCATCTGGATGAGAGAAA 1107
QY 229 GlyIleuLeuGlnAlaIleAlaArgGluGlnGluGluIlyLeuGlnArgValArgIly 248
    |||:::|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1108 AGTGAGTGTCTGACGCGGATTCACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1167
QY 249 LeuIleArgGlnIlyrGlyAspHisLeuGluIlySerSerIysLeuValGluSerAlaIle 268
    |||:::|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1168 CTATCCAGACAGTACACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1227
QY 269 GlnSerMetIleGluIleProGlnMetAlaLeuIlyrLeuGlnGlnAlaIlyGluLeuIleAsn 288
    |||:::|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1228 CAGTCTCTGAGCAGAGCTGGGGGAGCCACCTTCCTCTGACTGCAAGCAATCATCAAA 1287
QY 289 LysValGlyIleMetSerIysValGluLeuAlaGlyIlyrProGluIlyrProGlyIlyrGluSer 308
    |||:::|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1288 AGCATTTGTGAGAGCTTCCAAAGGCTGCCAGCTGGGAGAGACAGACAGAGAGAGAGAGAG 1347
QY 309 MetGluGlnPheSerValSerValGluHisValAlaGluMetLeuArgThrIleAspPhe 328
    |||:::|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1348 ATGAGACTTCTTACTTCTTGAATTAGACACATACAGACAGCGCTGAGAGCCATTCGACTTT 1407
QY 329 GlnProGlyAlaAlaGlyAspGluGluAspAspMetAla-----LeuAsp 344
    |||:::|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1408 -----GGACAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1455
QY 345 GlyIleGluGlyAsnAlaGlyLeuGluGlu 354
    |||:::|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1456 GAGGAGAGAGTCCACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1485
RESULT 16
HSA276484      1756 bp      mRNA      linear      PRI 04-JAN-2002
LOCUS      HSA276484
DEFINITION      Human mRNA for muscle specific RING finger 2 (MURF2) protein (MURF2
gene).
ACCESSION      AJ276484
VERSION      AJ276484.1 GI:18073355
KEYWORDS      MURF2 protein; RING finger protein.
SOURCE      human.
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE      1 (bases 1 to 1756)
AUTHORS      Stanchi,F.
TITLE      Characterisation of MURF2, a new muscle-specific RING finger
protein of the RbC family that associates with microtubules
JOURNAL      Unpublished
REFERENCE      2 (bases 1 to 1756)
AUTHORS      Stanchi,F.
TITLE      Direct Submission
JOURNAL      Submitted (20-MAR-2000) Stanchi F., CRIPI Biotechnology Centre,
Universita di Padova, Via G. Colombo 3, Padova, 35121, ITALY
FEATURES
source
    1..1756
        /organism="Homo sapiens"
        /db_xref="taxon:9606"
        /chromosome="1"
        /map="p35.1-35.2"
        /clone="STRAIT03036"
        /tissue_type="skeletal muscle"
        /clone_lib="pEAK8"
        /dev_stage="adult"
        1..1756
        /gene="MURF2"
        1..282
        /gene="MURF2"
        /number=1
        124..1185
        /gene="MURF2"
        /codon_start=1
        /product="muscle specific RING finger 2 (MURF2)"
        /protein_id="CAC81706.1"
        /db_xref="GI:18073356"

```

/translocation"ADYKSSSLTODNDSMPNMLEKOLICPLTELEAFTEPVVLLIPLCOHNK  
 RKCANDIEAOPANPYWTSRSGSSVMGMRPCRCHEHEDINBAGVGLIOLNIAVENT  
 DYPQESSRPLONGSHMCRKHEBERIYICLTCTVPCMSKVEEITKACKEPTLA  
 SVYOGKRETLNNGISIMLWAGNRVOTIITOLEDSRVATNSHMOVEEESLCAEPTLA  
 LITDKRSELIORIDOEKRSFTALILIOYQOCDLSKRYVATLAIOSIDPEGAGTFL  
 LITAKLSISYBAGSCGLKTEGCFENNDFTLIDLEHIALDALRAIDFCIDDEBEETI  
 LEEBDOEBEESTGKEGEGHQ

BASE COUNT	466 a	398 c	516 g	376 t
exon	283. .455			
	/gene="MURF2"			
	/number=2			
exon	456. .624			
	/gene="MURF2"			
	/number=3			
exon	625. .720			
	/gene="MURF2"			
	/number=4			
exon	725. .954			
	/gene="MURF2"			
	/number=5			
exon	955. .977			
	/gene="MURF2"			
	/number=6			
exon	978. .1102			
	/gene="MURF2"			
	/number=7			
exon	1103. .1174			
	/gene="MURF2"			
	/number=8			
exon	1175. .1756			
	/gene="MURF2"			
	/number=9			
polyA_signal	1720. .1725			
	/gene="MURF2"			
polyA_site	1756			
	/gene="MURF2"			



Percent Similarity: 78.00% Conservative: 63  
 Best Local Similarity: 60.00% Mismatches: 66  
 Query Match: 57.45% Indels: 11  
 DB: 9 Gaps: 4

US-09-908-988B-2 (1-366) x AK056942 (1-1764)

Oy 10 LeuLeuGlyAspAlaHisAsnMetAspAsnLeuGluIleuGlySerProIleCys 29  
 |||:|||||  
 Db 155 CTGATCCAGATGGGAAATCCCATGGAGAACTTGAGAGAGCAGCTGATCTGCCATATCTGC 214  
 |||:|||||  
 Oy 30 LeuGluMetPheSerLysProValIleLeuProCysGlnHisAsnLeuCysArgLys 49  
 |||:|||||  
 Db 215 CTGGAGATGTTTACCAAGACAGTGTGATCTTGCCGTCACGACCAACCTGTGGCCGGAG 274  
 |||:|||||  
 Oy 50 CysAlaAsnAspValPheGlnAlaSerAsnProLeuTrpGlnSerArgLysThr 69  
 |||:|||||  
 Db 275 TGTGCCAATGACATCTTCCAGGCTGCAAAATCCCTACTGAGACACCGCGGAGCTCAGTG 334  
 |||:|||||  
 Oy 70 ValSerSerGlyArgPheArgCysProSerCysArgHisGluValLeuAspArg 89  
 |||:|||||  
 Db 335 TCCATGCTCGAGAGCCGTTTCCGTCGCCACCTGCGCCAGCAGAGTGTATGATGATGCT 394  
 |||:|||||  
 Oy 90 HisGlyValTyrGlyLeuGlnArgAsnLeuValGluAsnIleLeuAspIleTyrLys 109  
 |||:|||||  
 Db 395 CACGGAGTTTACGGCCGTCGAGAGAACCTGCTGTGAGAGAACATCATCATCTACAA 454  
 |||:|||||  
 Oy 110 GlnGlu---SerSerArgProLeuHisAlaLysAlaGluGlnHisLeuMetCysGlu 128  
 |||:|||||  
 Db 455 CAGGAGTGTCTCAGTGGCCGCTG---CAGAAAGGCGAGTACCCTGTGCAAGAG 508  
 |||:|||||  
 Oy 129 HisGluAspGluLysIleAsnIleTyrCysLeuSerCysGluValProThrCysSerLeu 148  
 |||:|||||  
 Db 509 CACGAGAGAGAGAAATCAACATCTACTGTCTCAGCTGTGAGGTGCCACCTGCTCCATG 568  
 |||:|||||  
 Oy 149 CysLysValPheGlyAlaHisLysAspCysGluValAlaProLeuProThrIleTyrLys 168  
 |||:|||||  
 Db 569 TCCAGAGTGTGGGATCCACAGAGCCCTGCGAGGTGGCCCATTTGCAAGTGTCTCCAG 628  
 |||:|||||  
 Oy 169 ArgGlnLysSerGlyLeuSerAspGlyIleAlaMetLeuValAlaGlyAsnAspArgVal 188  
 |||:|||||  
 Db 629 GCACAAAGACGTGAATGATATCTGATCTCATCTGTGTGGGGGGAAGACCGTGTG 688  
 |||:|||||  
 Oy 189 GlnAlaValIleThrGlnMetGluGluValCysGlnThrIleGluAspAsnSerArgArg 208  
 |||:|||||  
 Db 689 CAGACCATATCATCTAGCTGGAGAGATCCCGTCGAGTGTGACCAAGAGAACAGTACCAG 748  
 |||:|||||  
 Oy 209 GlnLysGlnLeuLeuAsnGlnArgPheGluThrLeuCysAlaValLeuGluGluArgLys 228  
 |||:|||||  
 Db 749 GTAAAGAGAGAGCTGAGCCAGAGATTGACACGTTATGCCATCTCGATGAGAGAGAA 808  
 |||:|||||  
 Oy 229 GlyGluLeuLeuGlnAlaLeuAlaArgGluGlnGluLysLeuGlnArgValArgLys 248  
 |||:|||||  
 Db 809 AGTGAGTGTCTGACAGGGATCAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 868  
 |||:|||||  
 Oy 249 LeuIleArgGlnTyrGlyAspHisLeuGluGlySerSerLysLeuValGluSerAlaIle 268  
 |||:|||||  
 Db 869 CTGATCCACACATACAG 928  
 |||:|||||  
 Oy 269 GlnSerMetGluGluProGlnMetAlaLeuTyrLeuGlnGlnAlaLysGluLeuIleAsn 288  
 |||:|||||  
 Db 929 CAGTCCCTGAG 988  
 |||:|||||  
 Oy 289 LysValGlyAlaMetSerLysValGluLeuAlaGlyArgProGluProGlyTyrGlySer 308  
 |||:|||||  
 Db 989 AGCATTTGTGAAGCTTCAAGGGCTGCCACCTGCGGAGAGACAGAGAGAGAGAGAGAG 1048  
 |||:|||||  
 Oy 309 MetGlnGluPheSerValSerValGluHisValAlaGluMetLeuArgThrIleAspPhe 328  
 |||:|||||  
 Db 1049 ATGGACTTCTTACTTATTTGATTTAGAGACATATGACAGAGAGAGAGAGAGAGAGAGAG 1108  
 |||:|||||  
 Oy 329 GlnProGlyAlaAlaGluLysGluGluAspAspAspMetAla-----LeuAsp 344  
 |||:|||||

Db 1109 -----GGGACGATGAGAGAGAGAGAGATTCATTGAGAGAGAGATCAGAA 1156  
 Oy 345 GlyGluGluGlyAsnAlaGlyLeuGlu 354  
 |||:|||||  
 Db 1157 GAGAGAGATGTCACAGAGAGAGAGAGAGAA 1186  
 |||:|||||  
 RESULT 19  
 AF353673  
 LOCUS Homo sapiens  
 DEFINITION Homo sapiens iris ring finger protein mRNA, complete cds.  
 ACCESSION AF353673  
 VERSION AF353673.1 GI:13785923  
 KEYWORDS  
 SOURCE Homo sapiens.  
 ORGANISM Homo sapiens.  
 REFERENCE 1 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 TITLE 1 (bases 1 to 1791)  
 AUTHORS Wistow, G.  
 JOURNAL IRF: A Novel Ring Finger Protein From Iris  
 REFERENCE 2 Unpublished  
 AUTHORS Wistow, G.  
 TITLE 2 (bases 1 to 1791)  
 JOURNAL Direct Submission  
 TITLE Submitted (27-FEB-2001) MSF, NEI, 6/331, NIH, Bethesda, MD 20892, USA  
 FEATURES  
 source Location/Qualifiers  
 1..1791  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /chromosome="1"  
 /map="1p33-p31"  
 /tissue\_type="iris"  
 120..1181  
 /note="IRF"  
 /codon\_start=1  
 /product="iris ring finger protein"  
 /protein\_id="AAK39519.1"  
 /db\_xref="GI:13785924"  
 /translation="MDYKSSLIDQNPENILEKILICPICLEMTPKPVVILPCQHNLC RKANDIFQANPYWTSRGSVSMGGRFCPTGRHEVIMDRHGVGLHRLNLVNIID KIROKSSPRLKGSHPMCKHEHDEKINICYLCEVPTCSMKVGLHKACEVAPIO SVFQOKTELNLCISMLVAGNDRVOTIITOLEDSRVRYKNSHVKESLSEKFDLLA ILDEKSEILORTIOPEKRTSFTEALITQVQEDLDKTKIVETVITQSDDEGATFL LTRKOLIKTSIVESKSGCOLKRTGEGFENMDFIIDLHIALRAIDFGTDEEEEFI EEDDDEEESTKEKEEGHO"  
 BASE COUNT 505 a 398 c 512 g 376 t  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 4.04e-80 Length: 1791  
 Score: 1098.50 Matches: 210  
 Percent Similarity: 78.00% Conservative: 63  
 Best Local Similarity: 60.00% Mismatches: 66  
 Query Match: 57.45% Indels: 11  
 DB: 9 Gaps: 4  
 US-09-908-988B-2 (1-366) x AF353673 (1-1791)  
 Oy 10 LeuLeuGlyAspAlaHisAsnMetAspAsnLeuGluIleuGlySerProIleCys 29  
 |||:|||||  
 Db 138 CTGATCCAGATGGGAAATCCCATGGAGAACTTGAGAGAGCAGCTATCTGCCCTATCTGC 197  
 |||:|||||  
 Oy 30 LeuGluMetPheSerLysProValIleLeuProCysGlnHisAsnLeuCysArgLys 49  
 |||:|||||  
 Db 198 CTGGAGATGTTTACCAAGACAGTGTGATCTTGCCGTCACACAAACCTGTGGCGGAG 257  
 |||:|||||  
 Oy 50 CysAlaAsnAspValPheGlnAlaSerAsnProLeuTrpGlnSerArgLysThr 69  
 |||:|||||  
 Db 258 TGTGCCAATGACATCTTCCAGGCTGCAAAATCCCTACTGAGACAGCCGGGAGCTCAGTG 317  
 |||:|||||  
 Oy 70 ValSerSerGlyArgPheArgCysProSerCysArgHisGluValLeuAspArg 89  
 |||:|||||

D	b		318	TCCATGCTGGAGGCCGTTTCGGTCCGTGCACCTGACCACCGAGGTGATCATGGATCGT	377
O	y		90	HicgllyValIYrGlYleuglnArGaNsleuValGIuSnlIElleAsPIeTyrlYs	109
D	b		378	CACGACTGTACGGCCTGCAGAGAACCCTGCTGGTGGAACATCATCTGCATCACAA	437
O	y		110	GlnGlu---SerSerArgProLeuHIsAlaLysAlagInGlnHISlemCysGluLu	128
D	b		438	CAGGAGTGCCTCACTCGGCCGCTG-----CAGAAGGCCACTCACCCCATTGTCCAAGGAG	491
O	y		129	HISgluSpGLySIleasnIIeTYrCYsleuSerCYsgluValProThCYsSerleu	148
D	b		492	CACGAAGATGGAATAATCAATCATCTACTGCTCAAGTGTGAGTGAGGCCACCTGCTCATG	551
O	y		149	CysLySValPheGlYAlaHISlysAsPCysGIuValAlAPoleProTHleTYrlYs	168
D	b		552	TGCAGAGTGTGGATGCCAACAGGCCCTGCAGAGGTGGCCCCATTGCAAGAGTCTTCAG	611
O	y		169	ArgInLYsserGIuSerAsPGlYLelAMetleUValAlagIYNasAsPaRgAl	188
D	b		612	GGACAAGAACTGAACGTAAATACGTATCTTCATGCTGGTGGGGGAAAGAACCTGTG	671
O	y		189	GlNAlaValIIeThrgImetGluGluValCYsgInThrIIegIYNasPnsSerARgY	208
D	b		672	CAGAACATCATCACTCAAGCTGAGAGATTCGCCGTGAGTACCAAGAAACAGTACACAG	731
O	y		209	GlNylSGInleuLeuAsnGlnArGPhcGIuThrCYsAlaValleUGluGluArGLys	228
D	b		732	GTAAGGAGAGAGCTGAGCCAGAGATTTGACACGGTTGTATGCCATCTCGATGAGAGAAA	791
O	y		229	GlYluLeuIeuGlnAlaLeuAlaRgSluGluGluGluYlsleuGlnArGValARgLY	248
D	b		792	AGTAGTGTGTCAGCGGAGATCACGAGGAGAGAGAGAAAAAGCTTATCATCAGGCC	851
O	y		249	LeuIIeaRgGInTYrGlYSPHisleuGluCYsleuSerLySleuValGluSerAlle	268
D	b		852	CTCATCCAGCAGTACAGAGACAGCTGGACAAGTCCAAAGAGCTGTGTGAATGCCATC	911
O	y		269	GlNserMetGluGluProGlnMetaLaLeuTYrleuGlnGlnAlaLYsGluLeuIIeasn	288
D	b		912	CAGTCCCTGCGACAGACCCGTGGGGAGACCTTCTTCCTTGATGCCAACATCACTCAA	971
O	y		289	LysValGIAlaMetSerLySValGIuLeuAlaGlyARProGluTYrtyrSuber	308
D	b		972	AGCATTTGTGAAGCTTCCAAAGGCTGCCAGCTGGGAGAGACAGACAGGCGTTTAGAAC	1031
O	y		309	MetGluGlnPseSerValGIuHISValAlagImetleUArGThIleaspHe	328
D	b		1032	ATGACACTTCTTACTTGTGATTTAGAGCACAATAGCAGACGCCCTGAGAGCCATTACTTT	1091
O	y		329	GlNProGIlYaIaAlaGlyASPluGluNASpasPasmetaIa-----DeuASP	344
D	b		1092	-----GGACAGCATGAGAGAGAGAGAGAAATTCATTGAAGAGAAAGATCAGGA	1139
O	y		345	GlYgluGluGlyASNAlaGlyLeuGluGlu	354
D	b		1140	GAGGAAGAGTCCACAGAGAGGAGAGAGAA	1169
<b>RESULT 20</b>					
<b>AB047601</b>					
<b>LOCUS</b>					
AB047601 1746 bp mRNA linear PRI 11-OCT-2001					
<b>DEFINITION</b>					
Macaca fascicularis brain cDNA, clone:QnpA-10466.					
<b>ACCESSION</b>					
AB047601.1 GI:9929936					
<b>VERSION</b>					
AB047601.1 GI:9929936					
<b>KEYWORDS</b>					
f1s (full insert sequence).					
<b>SOURCE</b>					
Macaca fascicularis adult male brain parietal lobe cDNA to mRNA,					
clone_1lb:mecaque brain cDNA library QnpA clone:QnpA-10466.					
<b>ORGANISM</b>					
Macaca fascicularis					
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
Mammalia; Primates; Catarrhini; Cercopithecoidea;					
Cercopithecinae; Macaca.					
<b>REFERENCE</b>					
1 Osada,N., Hida,M., Kusuda,J., Tanuma,R., Iseki,K., Hirata,M.,					
<b>AUTHORS</b>					

TITLE	Suto, Y., Hirai, M., Terao, K., Suzuki, Y., Sugano, S. and Hashimoto, K.
JOURNAL	Assignment of 118 novel cDNAs of cynomolgus monkey brain to human chromosomes
MEDLINE	Gene 275 (1), 31-37 (2001)
AUTHORS	2. (bases 1 to 1746)
TITLE	Hashimoto, K., Oseada, N., Hida, M., Kusuda, J. and Sugano, S.
JOURNAL	Submitted (23-AUG-2000) Katsuyuki Hashimoto, National Institute of Infectious Diseases, Division of Genetic Resources; 23-1, Toyama 1-chome, Shinjuku-Ku, Tokyo 162-8640, Japan (E-mail:hashimuh.go.jp, URL:http://www.nih.go.jp/yoken/genebank/. Tel:81-3-5285-1111(ex.2120), Fax:81-3-5285-1181) Lab host: TOP10
COMMENT	Vector: pME18S-FU3 (Acc.No. AB009864) R. Site1: DraIII (CACTGTGTC) R. Site2: DraIII (CACCATGTC) Description: 1st strand cDNA was primed with an oligo(dT) primer [ATGCGCCCTTTTTTTTTTTTTTT]: double-stranded cDNA was synthesized using specific 5' and 3' primers and amplified by PCR. The PCR product was digested with SfiI and size selection was performed to exclude fragments <1.5kb. The SfiI-digested PCR product was cloned into distinct DraIII sites of pME18S-FU3. XhoI sites just outside the DraII sites can be used to isolate the cDNA insert. Libraries were constructed by Sugano et al.(University of Tokyo, Institute of Medical Science). Custom primer used for sequencing ( 5' end primer [CTTCTGCTTAAGAAGCTGG]; 3' end primer [CGACCTGCAGCCTCGACACA] ). Location/Qualifiers 1..1746 /organism="Macaca fascicularis" /db_xref="taxon:9541" /clone="Qnpa-10456" /sex="male" /tissue_type="brain parietal lobe" /clone_lib="macaque brain cDNA library Qnpa" /der_stage="adult" 567..1244 /codon_start=1 /product="hypothetical protein" /protein_id="PABI2125_1" /db_xref="GI:9929937" /translation="MSASLNKTSKSKSQQTMDNEKOLICPICEMETKPVILPCOHNLCKASDIFQASNPLYLPTFGITVMSGGFRPSCHEHYVDHREYGLQRULVBNIIDIQOESLRPKSDQPMCEHEERINITYLCNVEPVCSCXVFGARDDQVAALTHFQORSELSIDGIAIIVGSNDNRVOGVSLDEDTCKTEECRKOKOEICEKEFYDIYYSAGSGGO"
FEATURES	
SOURCE	
CDS	
BASE COUNT	523 a 396 c 387 g 440 t
ALIGNMENT SCORES:	
PRED. NO.:	3.03e-61 Length: 1746
Score:	866.00 Matches: 166
Percent Similarity:	77.25% Conservative: 31
Best Local Similarity:	65.10% Mismatches: 44
Query Match:	45.29% Indels: 14
DB:	Gaps: 9 2
US-09-908-988B-2 (1-366) x AB047601 (1-1746)	
OY	1 MetasnphethrValGlyPhenylsProLeuleugLiysAPAlAhIsAmetspasnleu 20
Db	567 ATGACGCGATCTGTCAAAAGTCATTTCCTTCACAAAGCAGACACACATGATTAACCTTA 626
OY	21 GluysglnleullelcyssProillecysteuenglumetPreserlysProValilleu 40
Db	627 GAGAAGCAACTATGTGCCATGCGCTTAGAGAGTGTCACGAACCCGCGTCAATTCTC 686
OY	41 ProCysglnhisAsnleucysArqscysAlaaasnsPvalPhgaIAlaserasnPro 60
Db	687 CCGTTCAGACAAACCTGTGTAGGAATGTGCCAATGATATTTTCAGAGCCCTTAACCG 746
OY	61 LeutrpGINserArgglySerthrThrValSerSerglyglYatrgPheargCysProSer 80

```
Db 747 TATTGCCCAAGAGAGACACCACCATGCAATGACAGGGCCGATTCGGCCCATCC 806
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY 81 CysArgHisGluValLeuAspArgHisGlyValTyrGlyLeuGlnArgAsnLeu 100
      |||
Db 807 TGTAGACATGAAGTGGTTGGATGACATGGGGTATATGACTTTCAGAGAACCTGCTG 866
      |||
QY 101 ValGluAsnIleIleAspIleTyrGlyGlnGluSerArgProLeuHisAlaLysAla 120
      |||
Db 867 GTGGAAAAATATCATTTACATCTTACAAAGAGAGAGAGCATCAACATCTACTGTGTAAC 920
      |||
QY 121 GluGlnHisLeuMetCysGluGlnHisGluAspGluLysIleAsnIleTyrCysLeuSer 140
      |||
Db 921 TCCGACACAGCCCATGTGTGAGAGCATGAGAGAGAGCATCAACATCTACTGTGTAAC 980
      |||
QY 141 CysGluValProThrCysSerLeuCysLysValPheGlyAlaHisLysAspCysGluVal 160
      |||
Db 981 TCGGAGTCCCACTGCTCTCTGTCAGAGGTGTTGGTGCACACAAGACTGCCAGGTG 1040
      |||
QY 161 AlaProLeuProThrIleTyrLysArgGlnLysSerGluLeuSerAspGlyIleAlaMet 180
      |||
Db 1041 GCTCCCTTACTCATGTGTTCAGAGACAGAAAGTCTGAGCTCAGCATGCGCATGCCATT 1100
      |||
QY 181 LeuValAlaGlyAsnAspArgValGlnAlaValIleThrGlnMetGluGluValCysGln 200
      |||
Db 1101 CTGGTGGGCGAGCAGCATGCGAGTCCAGAGGTGATCAGCCAGTGAAGACACCTGCAAA 1160
      |||
QY 201 ThrIleGluAspAsnSerArgArgGlnLysGlnLeuLeuAsnGlnArgPheGluThrLeu 220
      |||
Db 1161 ACCATTGAGGAATGTTGACAGAAACAGAAACAGAGCTTGTGTGAGAGTTT----- 1211
      |||
QY 221 CysAlaValLeuGluGlnLysGlyGlnLeuLeuGlnAlaLeuAlaArgGluGlnGlu 240
      |||
Db 1212 -----GATTACCTGTACGCGCACTGCAGGTGCTCAGTGA 1244
      |||
QY 241 GluLysLeuGlnArgValArgGlyLeuIleArgGlnTyrGlyAsp 255
      |||
Db 1245 AATTCAGTGTCAACTCAGAAAGTCAATGATTATCAATACATGAGAC 1289
      |||
```

Search completed: December 3, 2002, 15:33:45  
Job time : 2752 secs

DEFINITION BB140247 RIKEN full-length enriched, adult male bone Mus musculus  
 cDNA clone 9830168H24 3', mRNA sequence.  
 ACCESSION BB140247  
 VERSION BB140247.1 GI:8795184  
 KEYWORDS EST.  
 SOURCE house mouse.  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 REFERENCE 1 (bases 1 to 316)  
 AUTHORS Konno, H., Aizawa, K., Akahira, S., Akiyama, J., Arakawa, T., Carninci  
 P., Endo, T., Fukuda, S., Fukunishi, Y., Hara, A., Hayatsu, N.,  
 Hirozane, T., Hori, F., Ishii, Y., Ishikawa, J., Ishikawa, T., Itoh, M.,  
 Izawa, M., Kadota, K., Kagawa, I., Kai, C., Kawai, J., Kikuchi, N.,  
 Kiyosawa, H., Kojima, Y., Kondo, S., Koya, S., Kurihara, C., Kusakabe, M.,  
 Matsuyama, T., Miki, R., Mizuno, Y., Nakamura, M., Oda, H., Okazaki, Y.,  
 Ono, T., Owa, C., Saito, H., Sakai, C., Sato, K., Shibata, K., Shibata  
 Y., Shigemoto, Y., Shinagawa, A., Shiraki, T., Sogabe, Y., Sugahara, Y.,  
 Suzuki, H., Suzuki, H., Tagawa, A., Takahashi, F., Tominaga, N., Toya  
 T., Tsunoda, Y., Watahiki, A., Watanabe, S., Yamamura, T., Yamanaka, I.,  
 Yano, R., Yasunishi, A., Yokota, T., Yoshida, K., Yoshiki, A., Yoshino  
 M., Muramatsu, M. and Hayashizaki, Y.  
 TITLE RIKEN Mouse ESTs (Konno, H., et al.)  
 JOURNAL Unpublished (2000)  
 COMMENT Contact: Yoshihide Hayashizaki  
 Laboratory for Genome Exploration Research Group, RIKEN Genomic  
 Sciences Center (GSC), Yokohama Institute  
 The Institute of Physical and Chemical Research (RIKEN)  
 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan  
 Tel: 81-45-503-9222  
 Fax: 81-45-503-9216  
 Email: genome-res@gsc.riken.go.jp,  
 URL: http://genome.gsc.riken.go.jp/  
 FEATURES  
 source Location/Qualifiers  
 1. 316  
 /organism="Mus musculus"  
 /db\_xref="taxon:10090"  
 /clone="9830168H24"  
 /clone\_lib="RIKEN full-length enriched, adult male bone"  
 /sex="male"  
 /tissue\_type="bone"  
 /dev\_stage="adult"  
 /lab\_host="DH10B"  
 /note="Site\_1: SalI; Site\_2: BamHI; cDNA library was  
 prepared and sequenced in Mouse Genome Encyclopedia  
 Project of Genome Exploration Research Group in Riken  
 Genomic Sciences Center and Genome Science Laboratory in  
 RIKEN. Division of Experimental Animal Research in Riken  
 contributed to prepare mouse tissues. 1st strand cDNA was  
 primed with a primer [5'  
 GAGAGAGAGAAGGATCCAAGAGCTCTTTTTTTTTTTTTTTT 3'], cDNA was  
 prepared by using trehalose thermo-activated reverse  
 transcriptase and subsequently enriched for full-length by  
 cap-trapper. cDNA went through one round of normalization  
 to Rot = 10.0 and subtraction to Rot = 185.0. Second  
 strand cDNA was prepared with the primer adapter of  
 sequence [5' GAGAGAGAGATTCTCGAGTTAATTAAATTAATCCCCCCCCCCCC  
 3']. cDNA was cleaved with XhoI and BamHI. Vector: a  
 modified pBluescript KS(+) after bulk excision from Lambda  
 FLC I."  
 BASE COUNT 67 a 91 c 105 g 53 t  
 ORIGIN  
 Query Match 7.5%; Score 108; DB 10; Length 316;  
 Best Local Similarity 99.4%; Pred. No. 1.3e-40;  
 Matches 158; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1198 GGGGATGAAGAGGATGACGACATGGCTTTGGATGGGGAGGAGGGCAATGCGGGGCTGGAG 1257  
 Db 81 GGGGATGAAGAGGATGACGACATGGCTTTGGATGGGGAGGAGGGCAATGCGGGGCTGGAG 140  
 QY 1258 GAGGAGCGGCTGGACGTGCCAGAAGGCTCAGGCCTGCACTGACCCGACTCTGATCCAGAG 1317  
 Db 141 GAGGAGCGGCTGGACGTCCCAGAAGGCTCAGGCCTGCACTGACCCGACTCTGATCCAGAG 200  
 QY 1318 CGCACACCCGAAGCGGGAGCCAAGGGATGCTGAGGATCT 1356  
 Db 201 CGCACACCCGAAGCGGGAGCCAAGGGATGCTGAGGATCT 239

